

Boyle, Bryan (2012) *The structural chemistry of fluorinated nucleic acid components and possible implications for fluorinated DNA self-assembly*. PhD thesis.

<http://theses.gla.ac.uk/3198/>

Copyright and moral rights for this thesis are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

The structural chemistry of fluorinated nucleic acid components and possible implications for fluorinated DNA self-assembly

Bryan Boyle

Doctor of Philosophy Degree in Chemistry

School of Chemistry
University of Glasgow



September 2011

Supervisor: Professor Chick C. Wilson

Declaration

The thesis has been written in accordance with the University regulations and all work presented is original and performed by the author unless otherwise stated and referenced in the text.

Bryan Boyle, February 2012

Bryan Boyle
February 2012

Dedicated to Dr Jennifer S. Mathieson and Mrs Catherine Boyle

Acknowledgements

I would like to begin by thanking my supervisor Chick for all the help during my four years with the group. I was made to feel most welcome in the group and am very appreciative of the opportunity given to me in the group. Chick has been a fantastic boss to work for and a great person to get to know over the last four years.

I would also like to thank Lynne for the work she has put in to get me to this stage. I honestly believe without the work Lynne has put in this thesis may not have been possible. Lynne was also a great person to have round the group to keep us motivated and focused on the task at hand and also a good friend during my PhD. I do not envy Lynne for all the reading of my thesis I put her through. Thank you.

I would also like to thank the other members of the group that have made working on my PhD a very enjoyable experience. The two Craigs (big and wee) who have been a great laugh around the office and always there to talk about the football even when I do not want to. Thanks to Alan for his many discussions on television programs and computer games. Ioanna for the entertainment she has given in the upstairs office and Andras who was missed once he moved to Bath.

Thanks to all the other Chicklets that I have forgotten. All the people that have been associated with the group have been a pleasure to work with.

I would like to thank my parents for the support they have given me through my PhD. They have both been amazing and always been there for me when required. I cannot thank them enough. Paul and Louise who have been very supportive to me throughout the last four years and I love them to bits. Thanks also to the A-man for his entertainment at the football and pub before hand.

Last but certainly not least I would like to thank Jennifer. I was apprehensive about moving to Glasgow University at the end of my undergraduate degree but this spell at Glasgow has been amazing as I met Jennifer. She has been amazing support and the perfect person to have shared this experience with. Thank you so much.

Abstract

The main focus of this project has been the investigations into the effect that fluorine substitution has on structures of molecular complexes containing DNA base molecules. Previous work done on these molecules has been targeted at producing anti-cancer or anti-viral treatments, which reflects an important ultimate aim of investigations in this area. Success in this structural project would make a contribution towards this aim, by helping the understanding of the effects of fluorination on some of the interactions that control DNA self-assembly, notably base-pairing and stacking interactions, together with additional interactions involving fluorine. The specific aim of this project was to grow crystals of molecular complexes of cytosine and 5-fluorocytosine with co-molecules and use the structural descriptions to assess the difference between the structures and the influence the presence of the fluorine atom has on the structure when compared to the non-fluorinated equivalent.

Single crystal X-ray diffraction has been the main method of analysis, backed up by related techniques. Many crystallisations have been set up with cytosine and 5-fluorocytosine with a wide range of targeted co-molecules. 5-fluorouracil and uracil have also been used in related co-crystallisations with the aim of producing related complexes of these materials with the same co-molecules as used with cytosine and 5-fluorocytosine. In both families of complexes it was hoped that the fluorine would have an effect on the base-pairing motifs adopted in the structures, with the control being the non-fluorinated structure.

The program dSNAP has been used to give a comparison of the structures produced in this work with those already reported in the CSD. This program has given an indication of the classification of the standard primary bonding motifs and to analyse the other features commonly seen in base pairs such as buckling and propeller twisting.

The comprehensive series of complexes produced have provided the opportunity to examine significant structural trends, notable in the adoption of various base-pair motifs based on the Watson-Crick, Hoogsten and derived base-pairing patterns. The degree of proton transfer to the ring nitrogen of the cytosine could be rationalised in terms of the ΔpK_a values, and the proton transfer has a substantial effect on the base-pair motifs able to be adopted. The classification of the complex structures in terms of hetero (pseudo)-base-pairing has also been found to be of value. Extensive base-stacking and weaker interactions involving fluorine are also present; these have been analysed and found to have significant effects on the structures adopted, and hence may have future implications for the assembly of fluorinated DNA.

Table of Contents

Declaration	2
Acknowledgements	3
Abstract	4
Table of Contents	5
Table of Figures	10
Table of tables	36
Chapter 1 – Introduction	40
1.1 DNA	40
1.2 Nucleosides	43
1.3 Fluorinated Nucleosides.....	44
1.4 Crystal Engineering.....	45
1.5 Hydrogen Bonding	46
1.5.1 Strong Hydrogen Bonds.....	48
1.5.2 Moderate Hydrogen Bonds	48
1.5.3 Weak Hydrogen Bonds	48
1.5.3.1 Fluorine Hydrogen Bonds.....	49
1.5.4 Bifurcated Hydrogen Bonds.....	50
1.5.5 Charge Assisted Hydrogen Bonds	50
1.5.6 Proton Disorder	51
1.6 Hydrogen Transfer and pK _a matching	53
1.7 Halogen Bonding and Fluorine Interactions.	54
1.8 Other Intermolecular Interactions	55
1.8.1 π - interactions.....	55
1.8.2 Van der Waals	57
1.9 Nucleobases.....	58
1.9.1 Cytosine	58
1.9.2 5-Fluorocytosine	60
1.9.3 Thymine	63
1.9.4 Adenine	64
1.9.5 Guanine	66
1.3.6 Uracil.....	69
1.9.7 5-Fluorouracil.....	70
1.10. Classification of Base Pairs.....	73
1.10.1 Pseudo-Watson-Crick Base Pair.	73
1.10.2.Pseudo-Hoogsteen Base Pairing	74
1.10.2.1 Type A.....	74
1.10.2.2 Type B.....	74
1.10.2.3. Type C.....	75
1.10.3. Hydrogen Bonded Hetero-pseudo-dimer	75
1.10.4 No Primary Base Pairing Motif	76
1.11 Conformational parameters.....	76
Chapter 2 – Techniques.....	79
2.1 Introduction to Crystallography	79
2.2 Diffraction and Structure Solution	82
2.2.1 X-ray Scattering	82
2.2.2 Diffraction	83
2.2.3 Bragg's Law	84
2.2.4 Miller Indices	84
2.2.5 Reciprocal Lattice	85
2.2.6 Structure Factor	85
2.2.7. Structure Solution	87

2.2.7.1 The Phase Problem.....	87
2.2.7.2 Direct Methods for Structure Determination	87
2.2.7.3 Patterson Methods.....	88
2.2.8 Difference Fourier Synthesis.....	88
2.2.9 Structure Refinement	89
2.3 Techniques and Instrumentation	89
2.3.1 Single Crystal X-ray Diffraction.....	92
2.3.1.1 Selection of a Single Crystal	92
2.3.1.2 Data Collection.....	93
2.3.2 Powder X-ray Diffraction	94
2.3.2.1 Experimental Procedure	94
2.3.2.2 Phase Identification.....	95
2.3.2.3 Quantitative Analysis	95
2.3.2.4 Intensity Information.....	96
2.4 Crystallisation	96
2.5.1 Crystallisation by Slow Evaporation.....	96
2.5.2 Solvent Diffusion	97
2.5.3 Vapour Diffusion	97
2.5.4 Variable Temperature Parallel Crystallisation –The Microvate	98
2.6 Differential Scanning Calorimetry	99
2.7 Thermogravimetric Analysis (TGA).....	100
2.8 dSNAP	100
Chapter 3. Molecular Complexes of Cytosine with Benzoic Acid and Its Fluorinated Derivatives	105
3.1. 1:1 Molecular Complex of Cytosine and Benzoic acid [128].....	105
3.1.1 Benzoic Acid Bond Distances.....	107
3.1.2 Cytosine Bond Distances	108
3.2. 1:1:0.5 Molecular Complex of Cytosine and 2-Fluorobenzoic acid Hemihydrate .	113
3.3 1:1 Molecular Complex of Cytosinium 3-Fluorobenzoate	119
3.4. 2:1:1 Molecular Complex of Cytosine and 4-Fluorobenzoic acid Hydrate	123
3.5. Comparison of the Series of Molecular Complexes of Cytosine with Fluorosubstituted Benzoic acids.	126
3.5.1. Hydrogen Transfer and the Effect on the Crystallisation Ratio of the Product	127
3.5.2. Structural Motifs of the Cytosine Molecules	129
3.5.3 Structural Motifs Involving the Benzoic Acid Molecules	132
3.5.4. Base-pair stacking and the influence of water	133
3.5.5 The Anomalous Structure of the Cytosine with 3-Fluorobenzoic Acid Molecular Complex	134
3.5.6 Stacking Interactions.....	135
3.6. Molecular Complexes of 5-Fluorocytosine with Benzoic acid and the Series of Fluorobenzoic Acids	137
3.6.1. 1:1 Molecular Complex of 5-Fluorocytosine and Benzoic acid	137
3.6.2. 1:1 Molecular Complex of 5-Fluorocytosine and 2-Fluorobenzoic acid.....	141
3.6.3. 1:1 Molecular Complex of 5-Fluorocytosine and 3-Fluorobenzoic acid Solvate	145
3.6.4. 1:1 Molecular Complex of 5-Fluorocytosine and 4-Fluorobenzoic acid Solvate	151
3.7 Comparison of the Molecular Complexes of 5-Fluorocytosine with Benzoic Acids and their Fluorinated Derivatives.....	158
3.7.1. Structural Comparison of Molecular Complexes Formed Between 5- Fluorocytosine and Benzoic Acid and its Mono-Substituted Fluoro Derivatives.....	158
3.7.2. Common Hydrogen Bonding Motifs	159

3.7.3. The role of the benzoic acid molecule and fluorine interactions	159
3.7.4. Isostructural molecular complexes of 3-fluorobenzoic and 4-fluorobenzoic acid with 5-fluorocytosine	162
3.7.5. Molecular Complex of Benzoic Acid and 5-Fluorocytosine	165
3.3 Conclusions	166
Chapter 4. Molecular Complexes of Cytosine and 5-Fluorocytosine with Hydroxybenzoic Acids.	168
4.1. 1:1 Molecular Complex of Cytosinium Salicylate	168
4.2. 1:1 Molecular Complex of 5-Fluorocytosinium Salicylate [134]	173
4.3. 1:1 Molecular Complex of Cytosinium 5-Fluorosaliclyate	177
Molecular Complex.....	181
4.4. 2:1:2 Molecular Complex of 5-Fluorocytosine and 5-Fluorosaliclyic Acid Hydrate	182
4.5. 2:1:2 Molecular Complex of Cytosine and 5-Fluorosaliclyic Acid Hydrate	187
4.6. 2:2:1 Molecular Complex of Cytosinium and 3-Hydroxybenzoate Hemihydrate ..	192
4.7. 2:1 Molecular Complex of 5-Fluorocytosine and 3-Hydroxybenzoic Acid	199
4.8. 2:3:2 Molecular Complex of Cytosine and 4-Hydroxybenzoic Acid Hydrate	202
4.9. 1:1 Molecular Complex of 5-Fluorocytosine and 4-Hydroxybenzoic Acid	207
4.10. Structural Comparison of Molecular Complexes Formed Between Cytosine and 5-Fluorocytosine with Mono-Substituted Hydroxybenzoic Acids	210
4.10.1. The Effect of Proton Transfer	210
4.10.2. Complexes Containing Neutral Molecules Only	211
4.10.2.1. Comparison Between the Molecular Complexes of 5-Fluorocytosine with Benzoic Acid and 3-Hydroxybenzoic Acid	211
4.10.2.2. The Contrasting 4-Hydroxybenzoic Acid 5-Fluorocytosine Molecular Complex	213
4.10.3. Molecular Complexes Containing Charged Species Only	214
4.10.4. Molecular Complexes Showing Mixed Charge Molecular Components	218
4.10.5. The Effect of the Fluorine	221
4.10.6. Base Stacking Interactions	222
Chapter 5 Molecular Complexes of Urea and Thiourea with Cytosine and Uracil and their Fluorinated Derivatives	223
5.1. 1:1 Molecular Complex of 5-Fluorouracil and Urea	223
5.2. 1:1 Molecular Complex of Uracil and Urea	228
5.3 1:1 Molecular Complex of 5-Fluorouracil and Thiourea	231
5.4 3:1:1 Molecular Complex of 5-Fluorocytosine Thiourea Hydrate	235
5.5 2:1 Molecular Complex of 5-Fluorocytosine and Urea	242
5.6 Comparison of Structures	245
5.6.1 Isostructural molecular complexes of 5-fluorouracil with urea and uracil with urea	246
5.6.2. The Effect of Sulphur on the Molecular Complexes of Urea and Thiourea with 5-Fluorouracil	249
5.6.3 Cytosine Based Molecular Complexes	251
5.7. Conclusions	252
Chapter 6 Molecular Complexes of Disubstituted Benzoic acids with Cytosine, Uracil and their Fluorinated Derivatives	254
6.1. Polymorphic 1:1 Molecular Complexes of 5-Fluorocytosinium 2,6-Dihydroxybenzoate	254
6.1.1 1:1 Molecular Complex of 5-Fluorocytosinium 2,6-Dihydroxybenzoate Form I	254
6.1.2 1:1 Molecular Complex of 5-Fluorocytosinium 2,6-Dihydroxybenzoate Form II	256

6.1.3 Structural Comparison of the Two Polymorphs of 5-Fluorocytosinium 2,6-Dihydroxybenzoate	259
6.2 2:1:1 Molecular Complex of Cytosine and 2,6-Dihydroxybenzoic acid Hydrate ...	259
6.3. Polymorphs of the 1:1:1 Molecular Complex of Cytosinium 2,4-Dihydroxybenzoate Hydrate.....	263
6.3.1 1:1:1 Molecular Complex of Cytosinium 2,4-Dihydroxybenzoate Hydrate Form I	263
6.3.2. 1:1:1 Molecular Complex of Cytosinium 2,4-Dihydroxybenzoate Hydrate Form II	266
6.4 2:2:5 Molecular Complex of 5-Fluorocytosinium 3,5-Dihydroxybenzoate Hydrate	269
6.5 2:2:1 Molecular Complex of Cytosinium 3,5-Dihydroxybenzoate Hemihydrate....	274
6.6. 1:4:5 Molecular Complex of 5-Fluorouracil and 3,5-Dihydroxybenzoic Acid Hydrate.....	278
6.7. 1:1 Molecular Complex of 5-Fluorocytosine and 3,5-Dinitrobenzoic Acid	283
6.8. Structural Trends in Molecular Complexes with of Cytosine and 5-Fluorocytosine Disubstituted Hydroxybenzoic Acids	287
6.8.1 Molecular Complexes with 2,4-Dihydroxybenzoic Acid	288
6.8.2 Molecular Complexes of 2,6-Dihydroxybenzoic Acid	288
6.8.3 Molecular Complexes of 3,5-Dihydroxybenzoic Acid	288
6.8.4. Hydrogen Bonding Motifs	289
Chapter 7 Additional Molecular Complexes of Cytosine	292
7.1. Molecular Complexes of Cytosine with the Isomers Isonicotinic Acid and Nicotinic Acid.....	292
7.1.1 2:1:2 Cytosine and Isonicotinic acid Hydrate	292
7.1.2 1:1:1 Cytosinium Nicotinate Monohydrate.....	296
7.2. Molecular Complexes of Cytosine with Substituted Benzoic Acids	300
7.2.1 2:1 Cytosine and 2-nitrobenzoic acid.....	300
7.2.2 2:1 Cytosine and 3-bromobenzoic acid.....	304
7.2.3 1:1:1 Cytosinium 2,4,6-trimethylbenzoate monohydrate.....	306
7.3 Conclusions	309
Chapter 8. Systematic Study of Hydrogen Bonding Base Pairing Motifs using DSNAP and the CSD	312
8.1 Introduction to dSNAP.....	312
8.2. Base Pairing	313
8.2.1 Propeller Twist.....	314
8.2.2 Buckle	314
8.2.3 Hydrogen Bonding	314
8.2.4 dSNAP Study of Base Pairing	315
8.3 dSNAP Study of Adenine and Thymine Base Pairing.....	315
8.3.1 Clustering of Adenine Thymine Base Pairs	317
8.3.1.1 Analysis of the Red, Yellow and Green Clusters.....	318
8.3.1.2 Analysis of the Cyan Cluster.....	320
8.3.1.3 Analysis of the blue, green striped and orange clusters	321
8.3.1.3 Analysis of the Pink Cluster.....	323
8.2.1.4 Analysis of Orange Striped Cluster.....	323
8.3.1.6 Analysis of Cyan Striped, Blue Striped and Purple Striped Clusters	325
8.4 dSNAP Study of Cytosine and Guanine Base Pairing	327
8.4.1 Construction of the Search	327
8.4.2 Analysis of the Red Cluster	329
8.4.3 Analysis of Remaining Clusters.....	330
8.5 CSD Survey of Cytosine Base Pairing.....	332
8.5.1 Analysis of the Hydrogen Bonding Patterns Found in the CSD.....	332

8.5.2.1 Protonated Cytosine Molecules Only	332
8.5.2.2. Mixed Charge Cytosine Molecular Complexes	334
8.5.2.3 Neutral Cytosine Molecules Only	335
8.5.3. Summary	336
9. Conclusions	339
9.1. Molecular Complexes of Cytosine and 5-Fluorocytosine.....	339
9.1.1 pK _a Matching	339
9.1.1.1 1:1 Ratio	339
9.1.1.2 Anhydrous Non 1:1 Ratio Molecular Complexes.....	340
9.1.1.3 Molecular Complexes Containing Solvent Molecules.....	341
9.1.2 Base-Pair Motifs	342
9.2 Hydrogen Transfer	342
9.2.1 Total Hydrogen Transfer	342
9.2.2 Partial Hydrogen Transfer	344
9.2.3 No Hydrogen Transfer	345
9.3 Fluorine interactions.....	346
9.3.1 Fluorobenzoic acid Molecular Complexes.....	346
9.3.2 Hydroxybenzoic acid Molecular Complexes	348
9.3.3 Disubstituted Benzoic Acid Molecular Complexes	349
9.3.4 Thiourea and Urea Molecular Complexes	349
9.3.5. Prevalence of Fluorine Interactions	350
9.4 5-Fluorouracil and Uracil Molecular Complexes	350
9.5 Stacking interactions	353
9.6 dSNAP	354
9.7 Further Work	355
10. References	356
11. Appendix	363

Table of Figures

Chapter 1

FIGURE 1.1 DOUBLE HELICAL STRUCTURE OF OF DNA SHOWING THE HYDROGEN BONDS BETWEEN BASES AND THE SUGAR PHOSPHATE BACKBONE. P REPRESENTS THE PHOSPHATE BACKBONE, S THE SUGAR UNITS, AND A, T, G, C REPRESENT THE FOUR NUCLEOBASES [9].	41
FIGURE 1.2. THE FOUR NUCLEOBASES FOUND IN DNA, SPLIT INTO TWO GROUPS, PURINES (ADENINE AND GUANINE) AND PYRIMIDINES (THYMINE AND CYTOSINE).	42
FIGURE 1.3. THE BASE PAIRS FOUND IN DNA – GUANINE HYDROGEN BONDING WITH CYTOSINE AND ADENINE HYDROGEN BONDING WITH THYMINE.	43
FIGURE 1.4. THE DIFFERENCE BETWEEN A NUCLEOTIDE (LEFT – CONTAINING A PHOSPHATE GROUP) AND A NUCLEOSIDE (RIGHT – WITH NO PHOSPHATE GROUP).	43
FIGURE 1.5. BIFURCATED HYDROGEN-BOND SEEN IN PICOLINE AND CHLORANILIC ACID COMPLEXES [30].	46
FIGURE 1.6. P:C:P AND P:C:C:P COMPLEXES SEEN IN PICOLINE AND CHLORANILIC ACID COMPLEXES [30].	46
FIGURE 1.7. SCHEMATIC OF A CONVENTIONAL HYDROGEN BOND, DEMONSTRATED FOR WATER.	47
FIGURE 1.8. WEAK STRENGTH F...H-C HYDROGEN BOND IN THE MOLECULAR COMPLEX OF N,N,N',N'-TETRAMETHYLETHYLENEDIAMMONIUM BIS(DIFLUORO(PENTAFLUOROPHENYL)METHANIDE) [40].	49
FIGURE 1.9. THE MOST COMMON TYPE OF BIFURCATED HYDROGEN BOND.	50
FIGURE 1.10. POTENTIAL ENERGY PROFILE FOR AN ATOM BONDED TO ANOTHER LOCATED AT THE ORIGIN OF THE PLOT.	51
FIGURE 1.11. THREE DIFFERENT TYPES OF HYDROGEN BOND POTENTIAL: A) UNSYMMETRICAL DOUBLE MINIMUM, B) SYMMETRICAL DOUBLE WELL, C) FLAT BOTTOM SINGLE WELL.	52
FIGURE 1.12. AN EXAMPLE OF A DISORDERED HYDROGEN ATOM WITHIN A HYDROGEN BOND [47].	52
FIGURE 1.13. HALOGEN-HALOGEN INTERACTIONS – TYPE 1 (LEFT) AND TYPE 2 (RIGHT)	54
FIGURE 1.14. F...F INTERACTION PRESENT IN THE STRUCTURE OF N-(2-FLUOROBENZOYL)-N'-(P-TOLYL)THIOUREA [58].	55
FIGURE 1.15. AN EXAMPLE OF AN AROMATIC STACKING INTERACTION.	56
FIGURE 1.16. STACKING INTERACTIONS (A) C-H- π STACKING INTERACTION (B) EDGE TO EDGE STACKING INTERACTION AND (C) FACE TO FACE STACKING INTERACTION [22].	57
FIGURE 1.17. THE MOLECULAR STRUCTURE OF CYTOSINE.	58
FIGURE 1.18. THE STRUCTURE OF CYTOSINE MONOHYDRATE SHOWING THE EXTENDED HOMO-BASE PAIRED CHAIN (LEFT) AND THE WATER LINKING THE LAYERS (RIGHT) [64].	58
FIGURE 1.19. THE STRUCTURE OF ANHYDROUS CYTOSINE SHOWING THE EXTENDED HOMO-BASE PAIRED CHAIN (LEFT) AND THE CHAINS CHANGING ORIENTATION TO LINK SEPARATE LAYERS (RIGHT).	59
FIGURE 1.20. ASYMMETRIC UNIT OF THE CRYSTAL STRUCTURE OF THE MOLECULAR COMPLEX OF CYTOSINE AND TRIMESIC ACID REPORTED BY THOMAS, SHOWING THE HYDROGEN BONDED RING BONDING MOTIF [76].	59

FIGURE 1.21. THE PSEUDO-WATSON-CRICK BONDING MOTIF EXHIBITED BY THE CYTOSINIUM-HYDROGEN MALEATE-CYTOSINE MOLECULAR COMPLEX [71].	60
FIGURE 1.22. THE MOLECULAR STRUCTURE OF 5-FLUOROCYTOSINE.	61
FIGURE 1.23. EXTENDED HOMO-BASE PAIRED CHAIN FOUND IN 5-FLUOROCYTOSINE MONOHYDRATE [79].	61
FIGURE 1.24. WATER MOLECULES LINKING SEPARATE LAYERS FOUND IN 5-FLUOROCYTOSINE MONOHYDRATE [79].	61
FIGURE 1.25. HOMO-BASE PAIRED CHAIN (LEFT) AND THE STACKING OF THE BASES (RIGHT) FOUND IN THE ANHYDROUS FORM OF 5-FLUOROCYTOSINE [80].	62
FIGURE 1.26. HYDROGEN BONDED RING FOUND IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND SALICYLIC ACID [81].	62
FIGURE 1.27. SINGLE CHAIN (LEFT) AND TWO LINKED CHAINS (RIGHT) FOUND IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND SALICYLIC ACID [81].	63
FIGURE 1.28. HOMO-BASE PAIRED CHAIN FOUND IN THYMINE MONOHYDRATE [82].	63
FIGURE 1.29. WATER MOLECULES LINKING SEPARATE LAYERS IN THYMINE MONOHYDRATE [82].	63
FIGURE 1.30. THE HYDROGEN BONDING INVOLVING THE ADENINE MOLECULE IN THE TRIHYDRATE STRUCTURE. NO HOMO-BASE PAIRING IS FOUND [85].	64
FIGURE 1.31. WATER MOLECULES LINKING SEPARATE LAYERS IN THE MOLECULAR STRUCTURE OF ADENINE HYDRATE [85].	64
FIGURE 1.32. HOMO-BASE PAIRED CHAIN IN MOLECULAR COMPLEX OF ADENINE WITH PHTHALIC ACID [86].	65
FIGURE 1.33. PHTHALATE MOLECULES TIED TO THE HOMO-BASE PAIRED CHAIN IN MOLECULAR STRUCTURE OF ADENINE WITH PHTHALIC ACID [86].	65
FIGURE 1.34. PHTHALATE MOLECULES LINKING SEPARATE HOMO-BASE PAIRED CHAINS IN MOLECULAR COMPLEX OF ADENINE WITH PHTHALIC ACID [86].	65
FIGURE 1.35. THE BASE PAIRING OF ADENINE MOLECULES INTO LAYERS [87].	66
FIGURE 1.36. GUANINE MOLECULES HOMO-BASE PAIRING TO FORM SHEETS IN GUANINE MONOHYDRATE [88].	67
FIGURE 1.37. TWO DIFFERENT HOMO-BASE PAIRING MOTIFS IN GUANINE MONOHYDRATE [88].	67
FIGURE 1.38 WATER MOLECULES FORMING A CHAIN THAT LINKS DIFFERENT LAYERS OF THE GUANINE SHEETS IN THE MOLECULAR COMPLEX OF GUANINE [88].	67
FIGURE 1.39. HOMO-BASE PAIR PRESENT IN THE MOLECULAR COMPLEX OF GUANINE PICRATE MONOHYDRATE [89].	68
FIGURE 1.40. PICRATE MOLECULES COMBINING WITH WATER MOLECULES TO TERMINATE THE GUANINE HOMO-BASE PAIRED CHAIN IN THE MOLECULAR COMPLEX OF GUANINE PICRATE MONOHYDRATE [89].	68
FIGURE 1.41. CRYSTAL PACKING IN THE MOLECULAR COMPLEX OF GUANINE PICRATE MONOHYDRATE [89].	68
FIGURE 1.42. HOMO BASE PAIR IN THE CRYSTAL STRUCTURE OF URACIL [90].	69
FIGURE 1.43. HOMO BASE CHAIN (LEFT) AND URACIL SHEETS (RIGHT) IN THE CRYSTAL STRUCTURE OF URACIL [90].	69

FIGURE 1.44. BASE STACKING INTERACTIONS BETWEEN LAYERS IN THE CRYSTAL STRUCTURE OF URACIL [90].	69
FIGURE 1.45. HYDROGEN BONDING MOTIF (LEFT) AND ALL HYDROGEN BONDING INVOLVING URACIL MOLECULE (RIGHT) IN THE MOLECULAR COMPLEX OF URACIL AND MELAMINE [91].	70
FIGURE 1.46. THREE-DIMENSIONAL PACKING IN THE MOLECULAR COMPLEX OF URACIL AND MELAMINE [91].	70
FIGURE 1.47. SHEETS OF 5-FLUOROURACIL (LEFT) PACKED INTO LAYERS (RIGHT) IN THE CRYSTAL STRUCTURE OF 5-FLUOROURACIL FORM I [94].	71
FIGURE 1.48. CHAINS OF 5-FLUOROURACIL IN THE STRUCTURE OF 5-FLUOROURACIL FORM II [95].	71
FIGURE 1.49. FLUORINE INTERACTIONS IN THE STRUCTURE OF 5-FLUOROURACIL FORM II [95].	71
FIGURE 1.50. BASE PAIRED CHAIN IN THE MOLECULAR COMPLEX OF CYTOSINE AND 5-FLUOROURACIL MONOHYDRATE [67].	72
FIGURE 1.51. HYDROGEN BONDING LINKING THE CHAINS IN THE MOLECULAR COMPLEX OF CYTOSINE AND 5-FLUOROURACIL MONOHYDRATE [67].	72
FIGURE 1.52. SHEETS OF 5-FLUOROURACIL AND CYTOSINE (LEFT) AND BASE STACKING (RIGHT) IN THE MOLECULAR COMPLEX OF CYTOSINE AND 5-FLUOROURACIL MONOHYDRATE [67].	72
FIGURE 1.53. AN EXAMPLE OF A PSEUDO-WATSON-CRICK BONDING MOTIF [101].	73
FIGURE 1.54. AN EXAMPLE OF A TYPICAL TYPE A BONDING MOTIF [99].	74
FIGURE 1.55. AN EXAMPLE OF A TYPICAL TYPE B BONDING MOTIF [100]	74
FIGURE 1.56. AN EXAMPLE OF AN UNCOMMON TYPE C BONDING MOTIF [99].	75
FIGURE 1.57. AN EXAMPLE OF A TYPICAL HYDROGEN BONDED HETERODIMER [100]	76

Chapter 2

FIGURE 2.1. LEFT, THE UNIT CELL AND RIGHT, THE COMBINATION OF MANY UNIT CELLS TO PRODUCE THE FULL THREE-DIMENSIONAL CRYSTAL.	80
FIGURE 2.2. THE SEVEN CRYSTAL SYSTEMS AND THE FOURTEEN BRAVAIS LATTICES. [175]	81
FIGURE 2.3. CONSTRUCTIVE INTERFERENCE (LEFT) AND DESTRUCTIVE INTERFERENCE (RIGHT).	82
FIGURE 2.4. A TYPICAL DIFFRACTION PATTERN.	83
FIGURE 2.5. A GRAPHICAL REPRESENTATION OF BRAGG'S LAW.	84
FIGURE 2.6. TWO TYPES OF SINGLE CRYSTAL X-RAY DIFFRACTOMETERS USED IN THIS WORK. LEFT, A BRUKER APEX II CCD DIFFRACTOMETER AND RIGHT, A RIGAKU R-AXIS/RAPID IMAGE PLATE DIFFRACTOMETER.	92
FIGURE 2.7. GONIOMETER ON WHICH A CRYSTAL IS MOUNTED AND WHICH ALLOWS CENTRING OF THE CRYSTAL IN THE X-RAY BEAM.	93
FIGURE 2.8. TYPICAL SINGLE CRYSTALS USED FOR DIFFRACTION; THESE ARE OF INSULIN.	96
FIGURE 2.9. SCHEMATIC OF SLOW EVAPORATION CRYSTALLISATION.	97
FIGURE 2.10. SCHEMATIC OF SOLVENT DIFFUSION CRYSTALLISATION.	97
FIGURE 2.11. SCHEMATIC OF VAPOUR DIFFUSION CRYSTALLISATION.	97
FIGURE 2.12. THE MICROVATE PARALLEL CRYSTALLISER.	98

FIGURE 2.13. A TYPICAL DSC THERMOGRAM SHOWING THE DIFFERENT TYPES OF THERMAL EVENTS.....	99
FIGURE 2.14. A TGA THERMOGRAM SHOWING A STEP AS WATER IS RELEASED FROM THE SAMPLE BEFORE THE DECOMPOSITION TEMPERATURE IS REACHED[126].	100
FIGURE 2.15. DIFLUOROALKENE USED IN ANALYSIS USING THE DSNAP PROGRAM [127]	102
FIGURE 2.16. DENDROGRAM FOR THE ANALYSIS OF ACID-AMIDE INTERACTIONS [127]......	103
FIGURE 2.17 MMDS PLOT SHOWING THE SEPARATE CLUSTERS AND HOW CLOSELY RELATED THEY ARE [127].	104

Chapter 3

FIGURE 3.1. THE ASYMMETRIC UNIT OF THE CYTOSINE BENZOIC ACID MOLECULAR COMPLEX SHOWING THE FOUR INDEPENDENT MOLECULES.....	105
FIGURE 3.2. THE PSEUDO-WATSON-CRICK BASE CYTOSINE-CYTOSINIUM PAIR FOUND IN THE 1:1 MOLECULAR COMPLEX OF BENZOIC ACID AND CYTOSINE.	106
FIGURE 3.3. CONVENTIONAL WATSON-CRICK BASE PAIR INVOLVING CYTOSINE AND GUANINE.	106
FIGURE 3.4. THE LABELLING SCHEME FOR THE CYTOSINE MOLECULES (LEFT) AND THE BENZOIC ACID MOLECULES (RIGHT) IN THE CYTOSINE AND BENZOIC ACID MOLECULAR COMPLEX.	107
FIGURE 3.5. CYTOSINE HOMO-BASE PAIRS FORMING INTO A CHAIN ALONG THE <i>B</i> -AXIS IN THE MOLECULAR COMPLEX OF CYTOSINE AND BENZOIC ACID.	110
FIGURE 3.6. BENZOATE MOLECULES CONNECTING THE LAYERED CYTOSINE CHAINS ALONG THE <i>B</i> -AXIS IN THE MOLECULAR COMPLEX OF CYTOSINE AND BENZOIC ACID.	110
FIGURE 3.7. TWISTING OF THE CARBOXYLATE GROUP IN THE BENZOATE MOLECULE IN THE MOLECULAR COMPLEX OF CYTOSINE AND BENZOIC ACID.	110
FIGURE 3.8. NEUTRAL BENZOIC ACID HYDROGEN BONDED TO THE DEPROTONATED BENZOIC ACID IN THE MOLECULAR COMPLEX OF CYTOSINE AND BENZOIC ACID (LEFT) AND THE BENZOIC ACID CHAIN LINKED VIA WEAK O...H-C HYDROGEN BONDS (RIGHT).	111
FIGURE 3.9. $\Pi \cdots \Pi$ STACKING INTERACTIONS IN THE MOLECULAR COMPLEX OF CYTOSINE AND BENZOIC ACID	111
FIGURE 3.10. LINK BETWEEN SEPARATE SIX MOLECULE UNITS VIA WEAK INTERACTIONS IN THE MOLECULAR COMPLEX OF CYTOSINE AND BENZOIC ACID. THE MOLECULE INVOLVED IN WEAK INTERACTIONS BETWEEN BOTH BLOCKS IS REPRESENTED BY GREEN.....	112
FIGURE 3.11. STACKING INTERACTIONS IN THE MOLECULAR COMPLEX OF CYTOSINE AND BENZOIC ACID BETWEEN BUILDING BLOCKS WITH COLOUR SCHEME: RED (CHARGED CYTOSINE), GREEN (NEUTRAL CYTOSINE), ORANGE (CHARGED BENZOIC ACID) AND BLUE (NEUTRAL BENZOIC ACID)	112
FIGURE 3.12. THE MAJOR (GREEN) AND MINOR (BLUE) COMPONENTS OF THE 2-FLUOROBENZOIC ACID MOLECULES IN THE MOLECULAR COMPLEX OF CYTOSINE AND 2-FLUOROBENZOIC ACID HEMIHYDRATE.	113

FIGURE 3.13. THE ASYMMETRIC UNIT SHOWING PROTON TRANSFER FROM ONE BENZOIC ACID MOLECULE TO ONE CYTOSINE MOLECULE IN THE MOLECULAR COMPLEX OF CYTOSINE AND 2-FLUOROBENZOIC ACID HEMIHYDRATE.	113
FIGURE 3.14. THE PSEUDO-WATSON-CRICK BASE PAIR OBSERVED IN THE 1:1 MOLECULAR COMPLEX OF CYTOSINE AND 2-FLUOROBENZOIC ACID HEMIHYDRATE.	114
FIGURE 3.15. THE LABELLING SCHEME FOR THE CYTOSINE MOLECULES (LEFT) AND THE BENZOIC ACID MOLECULES (RIGHT) IN THE MOLECULAR COMPLEX OF CYTOSINE AND 2-FLUOROBENZOIC ACID HEMIHYDRATE	114
FIGURE 3.16. BASE PAIRED CHAIN LINKED BY HYDROGEN BONDS MAKING A DIAMOND SHAPE IN THE CYTOSINE 2-FLUOROBENZOIC ACID HEMIHYDRATE MOLECULAR COMPLEX.	115
FIGURE 3.17. STEPPED CYTOSINE CHAIN IN THE CYTOSINE 2-FLUOROBENZOIC ACID HEMIHYDRATE MOLECULAR COMPLEX.	115
FIGURE 3.18. 2-FLUOROBENZOATE AND WATER MOLECULES TIED TO THE CYTOSINE CHAIN IN THE CYTOSINE 2-FLUOROBENZOIC ACID HEMIHYDRATE MOLECULAR COMPLEX.	116
FIGURE 3.19. RING SHAPE FORMED BETWEEN TWO WATER AND TWO 2-FLUOROBENZOATE MOLECULES IN THE 2-FLUOROBENZOIC ACID AND CYTOSINE MOLECULAR COMPLEX.	116
FIGURE 3.20. RING CONNECTING LAYERS IN THE MOLECULAR COMPLEX OF CYTOSINE AND 2-FLUOROBENZOIC ACID HEMIHYDRATE.....	117
FIGURE 3.21. FLUORINE INTERACTION INVOLVING THE MAJOR COMPONENT OF THE DISORDERED 2-FLUOROBENZOIC ACID MOLECULE IN THE MOLECULAR COMPLEX OF CYTOSINE AND 2-FLUOROBENZOIC ACID HEMIHYDRATE.	117
FIGURE 3.22. FLUORINE INTERACTION INVOLVING THE MINOR COMPONENT OF THE DISORDERED 2-FLUOROBENZOATE (LEFT) AND 2-FLUOROBENZOIC ACID (RIGHT) MOLECULES IN THE MOLECULAR COMPLEX OF CYTOSINE AND 2-FLUOROBENZOIC ACID HEMIHYDRATE.	118
FIGURE 3.23. F...H-N HYDROGEN BOND PRESENT BETWEEN THE FLUORINE OF THE 2-FLUOROBENZOIC ACID MOLECULE AND THE AMINE GROUP OF A PROTONATED CYTOSINE MOLECULE IN THE MOLECULAR COMPLEX OF CYTOSINE AND 2-FLUOROBENZOIC ACID HEMIHYDRATE	118
FIGURE 3.24 STACKING INTERACTIONS IN THE MOLECULAR COMPLEX OF CYTOSINE AND 2-FLUOROBENZOIC ACID HEMIHYDRATE. COLOUR SCHEME, GREEN REPRESENTS NEUTRAL CYTOSINE AND RED REPRESENTS THE PROTONATED CYTOSINE MOLECULE.	119
FIGURE 3.25. THE ASYMMETRIC UNIT FOR THE MOLECULAR COMPLEX OF CYTOSINIUM 3-FLUOROBENZOATE INCLUDING ATOM LABELLING.....	120
FIGURE 3.26. HYDROGEN BONDING INVOLVING THE HETERODIMER IN THE MOLECULAR COMPLEX OF CYTOSINIUM 3-FLUOROBENZOATE.	120
FIGURE 3.27. ALTERNATING ORIENTATION OF CYTOSINIUM MOLECULES IN THE CYTOSINIUM 3-FLUOROBENZOATE MOLECULAR COMPLEX.	121
FIGURE 3.28. HETERODIMERS CONNECTED THROUGH HYDROGEN BONDS IN THE MOLECULAR COMPLEX OF CYTOSINIUM 3-FLUOROBENZOATE. ...	121
FIGURE 3.29. TWO SEPARATE CHAINS (FIGURE 3.28) LINKED TOGETHER VIA A SINGLE HYDROGEN BOND IN THE CYTOSINIUM 3-FLUOROBENZOATE MOLECULAR COMPLEX.....	121

FIGURE 3.30. FLUORINE INTERACTIONS IN THE CYTOSINIUM 3-FLUOROBENZOATE MOLECULAR COMPLEX.....	122
FIGURE 3.31 STACKING INTERACTIONS IN THE CYTOSINIUM 3-FLUOROBENZOATE MOLECULAR COMPLEX. COLOUR SCHEME, GREEN REPRESENTS CYTOSINIUM AND BLUE REPRESENTS THE 3-FLUOROBENZOATE MOLECULES.....	122
FIGURE 3.32. THE PSEUDO-WATSON-CRICK CYTOSINE-CYTOSINIUM BASE PAIR OBSERVED IN THE MOLECULAR COMPLEX OF 4-FLUOROBENZOIC ACID AND CYTOSINE HYDRATE WITH NUMBERING SCHEME.	123
FIGURE 3.33. CYTOSINE HOMO-BASE PAIRED CHAIN FORMED IN THE MOLECULAR COMPLEX OF CYTOSINE AND 4-FLUOROBENZOIC ACID HYDRATE WITH HYDROGEN BONDED SQUARES LINKING PSEUDO-WATSON-CRICK-BASE PAIRS ALONG THE B-AXIS.	124
FIGURE 3.34. FORMATION OF LADDER STRUCTURE IN THE MOLECULAR COMPLEX OF CYTOSINE AND 4-FLUOROBENZOIC ACID WITH THE WATER MOLECULE ACTING AS A SPACER.	124
FIGURE 3.35. COMPARISON OF THE SIX MEMBERED BLOCK IN THE MOLECULAR COMPLEX OF CYTOSINE AND BENZOIC ACID (TOP) AND THE EIGHT MEMBERED BLOCK IN THE MOLECULAR COMPLEX OF CYTOSINE AND 4-FLUOROBENZOIC ACID HYDRATE (BOTTOM) (THE WATER MOLECULES ARE IN GREEN).....	125
FIGURE 3.36. FLUORINE INTERACTIONS IN THE MOLECULAR COMPLEX OF CYTOSINE AND 4-FLUOROBENZOIC ACID HYDRATE.....	125
FIGURE 3.37 STACKING INTERACTIONS IN THE MOLECULAR COMPLEX OF CYTOSINE AND 4-FLUOROBENZOIC ACID HYDRATE WITH COLOUR SCHEME: RED REPRESENTING CYTOSINIUM MOLECULES, GREEN REPRESENTING NEUTRAL CYTOSINE MOLECULES, BLUE REPRESENTING THE 4-FLUOROBENZOATE MOLECULES AND YELLOW REPRESENTING THE WATER MOLECULES.....	126
FIGURE 3.38. THE ASYMMETRIC UNITS AND PROTON TRANSFER IN (A) CYTOSINE AND BENZOIC ACID, (B) CYTOSINE AND 2-FLUOROBENZOIC ACID HYDRATE (C) CYTOSINIUM 3-FLUOROBENZOATE (D) CYTOSINE AND 4-FLUOROBENZOIC ACID HYDRATE.....	127
FIGURE 3.39. PSEUDO-WATSON-CRICK BONDING MOTIF FOUND IN THE MOLECULAR COMPLEXES OF LEFT, CYTOSINE AND BENZOIC ACID, MIDDLE, CYTOSINE AND 2-FLUOROBENZOIC ACID HYDRATE AND RIGHT, CYTOSINE AND 4-FLUOROBENZOIC ACID HYDRATE.....	129
FIGURE 3.40. NUMBERING SCHEME FOR THE PSEUDO WATSON CRICK BASE PAIR APPLICABLE TO ALL THREE STRUCTURES.....	129
FIGURE 3.41. CYTOSINE HOMO-BASE PAIRED CHAIN FOUND IN THE MOLECULAR COMPLEXES OF LEFT, CYTOSINE AND BENZOIC ACID, MIDDLE, CYTOSINE AND 2-FLUOROBENZOIC ACID HYDRATE AND RIGHT, CYTOSINE AND 4-FLUOROBENZOIC ACID HYDRATE.....	130
FIGURE 3.42. STAGGERED CHAIN IN THE MOLECULAR COMPLEX OF CYTOSINE AND 2-FLUOROBENZOIC ACID HYDRATE WITH COLOUR SCHEME: GREEN FOR CYTOSINE MOLECULES AND RED FOR 2-FLUOROBENZOIC ACID MOLECULES.....	130
FIGURE 3.43. COMPARING CHAINS AND TWISTING OF THE BASE PAIRS IN THE MOLECULAR COMPLEX OF CYTOSINE AND 2-FLUOROBENZOIC ACID (LEFT) AND CYTOSINE AND BENZOIC ACID (RIGHT).....	131
FIGURE 3.44. CHANGES IN THE MOLECULES TIED TO EITHER SIDE OF THE HOMO-BASE PAIRED CHAIN FOUND IN THE MOLECULAR COMPLEXES OF LEFT, CYTOSINE AND BENZOIC ACID, MIDDLE, CYTOSINE AND 2-	

FLUOROBENZOIC ACID HYDRATE AND RIGHT, CYTOSINE AND 4- FLUOROBENZOIC ACID HYDRATE.....	132
FIGURE 3.45. THE SIX MEMBERED BLOCK FOUND IN THE MOLECULAR COMPLEX OF CYTOSINE AND BENZOIC ACID WITH THE NEUTRAL BENZOIC ACID TIED TO THE BLOCK VIA A HYDROGEN BOND.	133
FIGURE 3.46. RING SHAPE HELPING TO LINK SEVERAL LAYERS TOGETHER VIA HYDROGEN BONDING IN THE 2-FLUOROBENZOIC ACID AND CYTOSINE MOLECULAR COMPLEX.	133
FIGURE 3.47. FORMATION OF LADDER STRUCTURE IN THE MOLECULAR COMPLEX OF CYTOSINE AND 4-FLUOROBENZOIC ACID WITH THE WATER MOLECULE ACTING AS A SPACER.	134
FIGURE 3.48. FORMATION OF THE HYDROGEN BONDED RING IN THE MOLECULAR COMPLEX OF 3-FLUOROBENZOIC ACID AND CYTOSINE..	134
FIGURE 3.49 BASE STACKING INTERACTIONS SEEN IN THE MOLECULAR COMPLEXES OF (A) CYTOSINE AND BENZOIC ACID, (B) CYTOSINE AND 2- FLUOROBENZOIC ACID HYDRATE AND (C) CYTOSINE AND 4- FLUOROBENZOIC ACID HYDRATE.....	135
FIGURE 3.50. THE ASYMMETRIC UNIT OF THE MOLECULAR COMPLEX OF 5- FLUOROCYTOSINE AND BENZOIC ACID.	137
FIGURE 3.51 ALTERNATING TYPE A (BLACK BOX) AND TYPE B (RED BOX) HYDROGEN BONDING MOTIFS CREATING AN EXTENDED 5- FLUOROCYTOSINE CHAIN IN THE MOLECULAR COMPLEX OF 5- FLUOROCYTOSINE AND BENZOIC ACID.	138
FIGURE 3.52. HYDROGEN BONDS FOUND IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND BENZOIC ACID.....	138
FIGURE 3.53. EXTENDED CYTOSINE CHAIN WITH BENZOIC ACID MOLECULES TIED TO THE CHAIN (LEFT) VIA A MODERATE STRENGTH [130] O-H...O HYDROGEN BOND IN THE MOLECULAR COMPLEX OF 5- FLUOROCYTOSINE AND BENZOIC ACID.	139
FIGURE 3.54. THE PLANAR NATURE OF THE CYTOSINE CHAINS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND BENZOIC ACID..	139
FIGURE 3.55. FLUORINE INTERACTIONS TO THE CYTOSINE CHAIN IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND BENZOIC ACID..	140
FIGURE 3.56. FLUORINE INTERACTIONS BETWEEN SEPARATE CHAINS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND BENZOIC ACID..	140
FIGURE 3.57. ORIENTATION OF CHAINS (RED) THAT ARE CONNECTED TO A SECOND CHAIN (BLACK) IN THE MOLECULAR COMPLEX OF 5- FLUOROCYTOSINE AND BENZOIC ACID.	140
FIGURE 3.58. STACKING INTERACTIONS (LEFT) AND CLOSE UP OF BASE STACKING INTERACTION (RIGHT) FOUND IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND BENZOIC ACID.....	141
FIGURE 3.59. ASYMMETRIC UNIT OF THE MOLECULAR COMPLEX OF 5- FLUOROCYTOSINE AND 2-FLUOROBENZOIC ACID.....	142
FIGURE 3.60. THE ALTERNATING TYPE A (RED) AND TYPE B (BLACK) BASE PAIRING IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE WITH 2- FLUOROBENZOIC ACID.....	142
FIGURE 3.61. HYDROGEN BONDS CONNECTING THE BENZOIC ACID TO THE 5- FLUOROCYTOSINE CHAIN IN THE MOLECULAR COMPLEX OF 5- FLUOROCYTOSINE AND 2-FLUOROBENZOIC ACID.....	143
FIGURE 3.62. EXTENDED CHAIN (LEFT) AND THE MAIN CONNECTION BETWEEN THE 2-FLUOROBENZOIC ACID AND THE 5-FLUOROCYTOSINE CHAIN (RIGHT) IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 2-FLUOROBENZOIC ACID.....	143

FIGURE 3.63. PLANARITY OF THE 5-FLUOROCYTOSINE CHAIN AND THE TWISTING OF THE 2-FLUOROBENZOIC ACID WITH RESPECT TO THIS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 2- FLUOROBENZOIC ACID.....	144
FIGURE 3.64. FLUORINE INTERACTIONS BETWEEN THE 5-FLUOROCYTOSINE CHAIN AND THE TETHERED 2-FLUOROBENZOIC ACID MOLECULES IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 2- FLUOROBENZOIC ACID.....	144
FIGURE 3.65. FLUORINE INTERACTIONS LINKING SEPARATE CHAINS (LEFT) AND THE ORIENTATION OF THE CHAINS CONNECTED VIA THESE FLUORINE INTERACTIONS (RIGHT) IN THE MOLECULAR COMPLEX OF 5- FLUOROCYTOSINE AND 2-FLUOROBENZOIC ACID.....	145
FIGURE 3.66. STACKING INTERACTIONS BETWEEN SEPARATE 5- FLUOROCYTOSINE HOMO-BASE PAIRED CHAINS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 2-FLUOROBENZOIC ACID.	145
FIGURE 3.67. THE ASYMMETRIC UNIT OF THE MOLECULAR COMPLEX OF 5- FLUOROCYTOSINE AND 3-FLUOROBENZOIC ACID.....	146
FIGURE 3.68. THE ALTERNATING TYPE A (BLACK) AND TYPE B (RED) BASE PAIRING IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE WITH 3- FLUOROBENZOIC ACID.....	146
FIGURE 3.69. EXTENDED HOMO-BASE PAIRED CHAIN WITH 3- FLUOROBENZOIC ACID MOLECULES CONNECTED TO THE 5- FLUOROCYTOSINE CHAIN IN THE MOLECULAR COMPLEX OF 5- FLUOROCYTOSINE AND 3-FLUOROBENZOIC ACID.....	146
FIGURE 3.70. FLUORINE INTERACTIONS LINKING SEPARATE CHAINS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 3- FLUOROBENZOIC ACID.....	147
FIGURE 3.71. SEPARATE CHAINS LINKED VIA HYDROGEN BONDS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 3- FLUOROBENZOIC ACID (GREEN REPRESENTS THE HYDROGEN BONDS THAT MAKE USE OF THE FLUORINE ATOMS AND THE RED AND BLACK REPRESENT SEPARATE CHAINS).....	148
FIGURE 3.72. TGA THERMOGRAM OF THE MOLECULAR COMPLEX OF 5- FLUOROCYTOSINE AND 3-FLUOROBENZOIC ACID.....	149
FIGURE 3.73. DSC THERMOGRAM OF THE MOLECULAR COMPLEX OF 5- FLUOROCYTOSINE AND 3-FLUOROBENZOIC ACID.....	150
FIGURE 3.74 SPACEFILL REPRESENTATION OF THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 3-FLUOROBENZOIC ACID SHOWING THE PRESENCE OF A SIGNIFICANT VOID AT REPEATING INTERVALS THROUGHOUT THE STRUCTURE.....	150
FIGURE 3.75. STACKING INTERACTIONS BETWEEN 5-FLUOROCYTOSINE CHAINS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 3- FLUOROBENZOIC ACID.....	151
FIGURE 3.76. STACKING INTERACTIONS BETWEEN 3-FLUOROBENZOIC ACID MOLECULES IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 3-FLUOROBENZOIC ACID.....	151
FIGURE 3.77. ASYMMETRIC UNIT OF THE MOLECULAR COMPLEX OF 5- FLUOROCYTOSINE AND 4-FLUOROBENZOIC ACID.....	152
FIGURE 3.78. THE ALTERNATING TYPE A AND TYPE B BASE PAIRING IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE WITH 4- FLUOROBENZOIC ACID.....	152
FIGURE 3.79. EXTENDED HOMO-BASE PAIRED CHAIN OF 5-FLUOROCYTOSINE WITH 4-FLUOROBENZOIC ACID MOLECULES CONNECTED TO THE CHAIN	

IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 4-FLUOROBENZOIC ACID.....	153
FIGURE 3.80. FLUORINE INTERACTIONS LINKING SEPARATE CHAINS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 4-FLUOROBENZOIC ACID.....	153
FIGURE 3.81. SEPARATE CHAINS LINKED VIA WEAK HYDROGEN BONDS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 4-FLUOROBENZOIC ACID (GREEN REPRESENTS THE HYDROGEN BONDS THAT MAKE USE OF THE FLUORINE ATOMS AND THE RED AND BLACK REPRESENT SEPARATE CHAINS).....	154
FIGURE 3.82. TGA THERMOGRAM OF THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 4-FLUOROBENZOIC ACID.....	155
FIGURE 3.83. DSC THERMOGRAM OF THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 4-FLUOROBENZOIC ACID.....	155
FIGURE 3.84. SPACEFILL REPRESENTATION OF THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 4-FLUOROBENZOIC ACID SHOWING THE SIGNIFICANT VOID.....	156
FIGURE 3.85. STACKING INTERACTIONS FOUND BETWEEN 5-FLUOROCYTOSINE CHAINS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 4-FLUOROBENZOIC ACID.....	156
FIGURE 3.86. STACKING INTERACTIONS FOUND BETWEEN 4-FLUOROBENZOIC ACID MOLECULES IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 4-FLUOROBENZOIC ACID.	156
FIGURE 3.87. TYPE A AND TYPE B HYDROGEN BONDING MOTIFS SEEN IN THE EXTENDED HOMO-BASE PAIRED CHAINS IN THE MOLECULAR COMPLEXES OF 5-FLUOROCYTOSINE WITH BENZOIC ACID AND 2-,3- AND 4-FLUOROBENZOIC ACID.....	159
FIGURE 3.88. FLUORINE INTERACTIONS LINKING SEPARATE CHAINS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 2-FLUOROBENZOIC ACID.....	160
FIGURE 3.89. FLUORINE INTERACTIONS LINKING SEPARATE CHAINS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 2-FLUOROBENZOIC ACID SHOWING EXTENDED CHAINS (ONE REPRESENTED BY BLUE AND ONE BY GREEN).....	160
FIGURE 3.90. FLUORINE INTERACTIONS BETWEEN THE LAYERS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 2-FLUOROBENZOIC ACID.....	161
FIGURE 3.91. EXTENDED CHAINS IN THE MOLECULAR COMPLEXES OF 3-FLUOROBENZOIC ACID AND 5-FLUOROCYTOSINE (LEFT) AND 4-FLUOROBENZOIC ACID AND 5-FLUOROCYTOSINE (RIGHT).....	162
FIGURE 3.92. EXTENDED CHAINS LINKED TOGETHER VIA FLUORINE INCLUDED BASE-PAIRED MOTIF PRESENT IN THE MOLECULAR COMPLEXES OF 3-FLUOROBENZOIC ACID AND 5-FLUOROCYTOSINE (LEFT) AND 4-FLUOROBENZOIC ACID AND 5-FLUOROCYTOSINE (RIGHT).....	162
FIGURE 3.93. CLOSE UP OF FLUORINE BASE PAIRED STYLE MOTIF LINKING SEPARATE CHAINS IN THE 3-FLUOROBENZOIC ACID AND 5-FLUOROCYTOSINE (LEFT) AND THE 4-FLUOROBENZOIC ACID AND 5-FLUOROCYTOSINE MOLECULAR COMPLEXES.	163
FIGURE 3.94. STRUCTURE OVERLAY OF THE 3-FLUOROBENZOIC ACID AND 5-FLUOROCYTOSINE MOLECULAR COMPLEX (GREEN) AND 4-FLUOROBENZOIC ACID AND 5-FLUOROCYTOSINE MOLECULAR COMPLEX (BLUE) SHOWING THE LINKS BETWEEN THE CHAINS.	163

FIGURE 3.95. STRUCTURE OVERLAY OF THE 3-FLUOROBENZOIC ACID AND 5-FLUOROCYTOSINE MOLECULAR COMPLEX (GREEN) AND 4-FLUOROBENZOIC ACID AND 5-FLUOROCYTOSINE MOLECULAR COMPLEX (BLUE).	163
FIGURE 3.96. SPACEFILL REPRESENTATION TO SHOW THE VOIDS PRESENT IN THE MOLECULAR COMPLEXES OF 3-FLUOROBENZOIC ACID AND 5-FLUOROCYTOSINE (LEFT) AND 4-FLUOROBENZOIC ACID AND 5-FLUOROCYTOSINE (RIGHT).	164
FIGURE 3.97. FLUORINE INTERACTIONS HELPING TO LINK SEPARATE CHAINS TOGETHER IN THE 5-FLUOROCYTOSINE AND BENZOIC ACID MOLECULAR COMPLEX.	165
FIGURE 3.98. STACKED LAYERS IN THE 5-FLUOROCYTOSINE AND BENZOIC ACID MOLECULAR COMPLEX.	166
FIGURE 3.99. CHAINS RUNNING PERPENDICULAR TO EACH OTHER IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND BENZOIC ACID (LEFT) AND THE STACKED LAYERS PRESENT IN THE 5-FLUOROCYTOSINE AND 2-FLUOROBENZOIC ACID MOLECULAR COMPLEX WITH NO VOID IN EITHER STRUCTURE.	166

Chapter 4

FIGURE 4.1. THE HETERODIMER FORMED IN THE CYTOSINIUM SALICYLATE MOLECULAR COMPLEX.	168
FIGURE 4.2. THE HYDROGEN BONDING ENVIRONMENT OF THE CYTOSINIUM ION IN THE CYTOSINIUM SALICYLATE MOLECULAR COMPLEX.	169
FIGURE 4.3. CHAINS OF HYDROGEN BONDED HETERODIMERS IN THE CYTOSINIUM SALICYLATE MOLECULAR COMPLEX.	170
FIGURE 4.4. LEFT, THE HERRINGBONE ARRANGEMENT OF HYDROGEN BONDED L SHAPED CHAINS. ALL L SHAPED UNITS ARE SYMMETRY RELATED; RIGHT, LINK BETWEEN SEPARATE CHAINS TO CREATE THE L SHAPE VIA MODERATE N-H...O HYDROGEN BONDS.	170
FIGURE 4.5. BASE STACKING IN THE CYTOSINIUM SALICYLATE MOLECULAR COMPLEX. THE UNITS ARE STAGGERED WITH RESPECT TO ONE ANOTHER AND THE PLANES ARE GENERATED THROUGH SALICYLATE MOLECULES.	170
FIGURE 4.6. THE HERRINGBONE ARRANGEMENT OF THE HYDROGEN BONDED L SHAPED CHAINS WITH WEAK HYDROGEN BONDS LINKING MOLECULES TO EITHER END OF THE CHAINS.	171
FIGURE 4.7. 5-FLUOROCYTOSINIUM SALICYLATE HYDROGEN BONDED HETERODIMER.	173
FIGURE 4.8. THE HYDROGEN BONDS INVOLVING A 5-FLUOROCYTOSINIUM MOLECULE IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINIUM SALICYLATE.	174
FIGURE 4.9. LEFT, CHAIN OF HYDROGEN BONDED HETERODIMERS IN THE 5-FLUOROCYTOSINIUM SALICYLATE MOLECULAR COMPLEX; RIGHT, THE HERRINGBONE STRUCTURE OF HYDROGEN BONDED L SHAPED CHAINS.	174
FIGURE 4.10. BASE-STACKING IN THE 5-FLUOROCYTOSINIUM SALICYLATE MOLECULAR COMPLEX.	175
FIGURE 4.11. MOLECULES HYDROGEN BONDED TO THE TIPS OF THE L-SHAPE CREATED BY THE PAIRS OF LEFT, CYTOSINIUM SALICYLATE AND RIGHT, 5-FLUOROCYTOSINIUM SALICYLATE MOLECULAR COMPLEXES.	176

FIGURE 4.12. BASE-STACKING INTERACTIONS AND FLUORINE INTERACTIONS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINIUM SALICYLATE. THE FLUORINE INTERACTIONS ARE REPRESENTED BY PURPLE DASHED LINES AND THE BASE STACKING INTERACTIONS BY LUMINOUS ORANGE DASHED LINES.....	176
FIGURE 4.13. INFLUENCE OF THE FLUORINE ATOM IN THE HYDROGEN BONDING IN THE RELATED STRUCTURES OF 5-FLUOROCYTOSINIUM SALICYLATE (LEFT) AND CYTOSINIUM SALICYLATE (RIGHT).....	177
FIGURE 4.14. HYDROGEN BONDED HETERODIMER FORMED IN THE CYTOSINIUM 5-FLUOROSALICYLATE MOLECULAR COMPLEX.....	178
FIGURE 4.15. LEFT. THE HYDROGEN BONDING ENVIRONMENT OF THE CYTOSINIUM MOLECULE IN THE CYTOSINIUM 5-FLUOROSALICYLATE MOLECULAR COMPLEX; RIGHT, CHAIN OF HYDROGEN-BONDED HETERODIMERS.	178
FIGURE 4.16. FLUORINE INTERACTIONS AROUND THE LINK BETWEEN EACH ARM OF THE L SHAPE UNIT IN THE CYTOSINIUM 5-FLUOROSALICYLATE MOLECULAR COMPLEX.....	179
FIGURE 4.17. THE HERRINGBONE STRUCTURE OF HYDROGEN BONDED L SHAPED CHAINS IN THE CYTOSINIUM 5-FLUOROSALICYLATE MOLECULAR COMPLEX.....	179
FIGURE 4.18. BASE-STACKING IN THE CYTOSINIUM 5-FLUOROSALICYLATE MOLECULAR COMPLEX.....	180
FIGURE 4.19. MOLECULES HYDROGEN BONDED TO THE TIPS OF THE L-SHAPE CREATED BY THE PAIRS IN THE MOLECULAR COMPLEXES OF (A), CYTOSINIUM SALICYLATE, (B) 5-FLUOROCYTOSINIUM SALICYLATE AND (C) CYTOSINIUM 5-FLUOROSALICYLATE.....	181
FIGURE 4.20. FLUORINE INTERACTIONS IN THE MOLECULAR COMPLEX OF CYTOSINIUM 5-FLUOROSALICYLATE.....	182
FIGURE 4.21. LEFT, THE ASYMMETRIC UNIT OF THE MOLECULAR COMPLEX SHOWING PROTON TRANSFER FROM ONE 5-FLUOROSALICYLIC MOLECULE TO ONE 5-FLUOROCYTOSINE MOLECULE; RIGHT THE PSEUDO WATSON-CRICK BASE PAIR.	182
FIGURE 4.22. LEFT, THE PSEUDO-WATSON-CRICK AND TYPE B HYDROGEN BONDING MOTIFS AND RIGHT, THE TERMINATION OF THE CHAIN IN THE 5-FLUOROCYTOSINE AND 5-FLUOROSALICYLIC ACID HYDRATE MOLECULAR COMPLEX TO CREATE AN EIGHT MOLECULE UNIT.....	183
FIGURE 4.23. THE HYDROGEN BONDS INVOLVING THE FOUR MOLECULE 5-FLUOROCYTOSINE 5-FLUOROCYTOSINIUM UNIT IN THE 5-FLUOROCYTOSINE 5-FLUOROSALICYLIC ACID HYDRATE MOLECULAR COMPLEX. THE WATER MOLECULES WHICH TERMINATE THE CYTOSINE CHAINS ARE SHOWN IN YELLOW AND THOSE THAT CONNECT PLANES IN DIFFERENT LAYERS ARE IN GREEN.....	185
FIGURE 4.24 THE SIX MOLECULE HYDROGEN BONDED RING IN THE 5-FLUOROCYTOSINE 5-FLUOROSALICYLIC ACID HYDRATE MOLECULAR COMPLEX CONNECTING TWO EIGHT MOLECULE UNITS TOGETHER.....	185
FIGURE 4.25. HYDROGEN BONDS INVOLVING THE WATER MOLECULE WHICH CONNECTS THE LAYERS TOGETHER IN THE 5-FLUOROCYTOSINE 5-FLUOROSALICYLIC ACID HYDRATE MOLECULAR COMPLEX.....	186
FIGURE 4.26. FLUORINE INTERACTIONS IN THE 5-FLUOROCYTOSINE 5-FLUOROSALICYLIC ACID HYDRATE MOLECULAR COMPLEX. LEFT, FLUORINE-FLUORINE INTERACTION BETWEEN 5-FLUOROCYTOSINIUM MOLECULES; RIGHT, WEAK C-H...F HYDROGEN BOND.....	186

FIGURE 4.27. STACKING INTERACTIONS IN THE 5-FLUOROCYTOSINE 5-FLUOROSALICYLIC ACID HYDRATE MOLECULAR COMPLEX.....	187
FIGURE 4.28. LEFT, THE ASYMMETRIC UNIT OF THE MOLECULAR COMPLEX OF CYTOSINE 5-FLUOROSALICYLIC ACID HYDRATE SHOWING PROTON TRANSFER FROM THE 5-FLUOROSALICYLIC ACID MOLECULE TO ONE CYTOSINE MOLECULE; RIGHT, THE PSEUDO WATSON-CRICK BASE PAIR.	188
FIGURE 4.29. LEFT, LINK BETWEEN TWO PSEUDO-WATSON CRICK BASE PAIRS VIA TYPE B HYDROGEN BONDING MOTIF; RIGHT, TERMINATION OF THE CHAIN IN THE CYTOSINE AND 5-FLUOROSALICYLIC ACID HYDRATE MOLECULAR COMPLEX.....	189
FIGURE 4.30. THE HYDROGEN BONDS INVOLVING THE FOUR MOLECULE CYTOSINE CYTOSINIUM UNIT IN THE CYTOSINE 5-FLUOROSALICYLIC ACID HYDRATE MOLECULAR COMPLEX. THE WATER MOLECULES WHICH TERMINATE THE CYTOSINE CHAINS ARE SHOWN IN GREEN AND THOSE THAT CONNECT PLANES IN DIFFERENT LAYERS ARE IN PURPLE.	190
FIGURE 4.31. SIX MOLECULE HYDROGEN BONDED RING IN THE CYTOSINE 5-FLUOROSALICYLIC ACID HYDRATE MOLECULAR COMPLEX.....	190
FIGURE 4.32. HYDROGEN BONDS INVOLVING THE WATER MOLECULE CONNECTING THE LAYERS IN THE CYTOSINE 5-FLUOROSALICYLIC ACID HYDRATE MOLECULAR COMPLEX.....	190
FIGURE 4.33. WEAK C-H...F HYDROGEN BONDS REPRESENTED BY A BLACK DASHED LINE, IN THE CYTOSINE 5-FLUOROSALICYLIC ACID HYDRATE MOLECULAR COMPLEX.....	191
FIGURE 4.34. STACKING INTERACTIONS PRESENT IN THE CYTOSINE 5-FLUOROSALICYLIC ACID HYDRATE MOLECULAR COMPLEX.....	191
FIGURE 4.35. LEFT, OVERLAY OF THE STRUCTURES OF CYTOSINE 5-FLUOROSALICYLIC ACID HYDRATE (WHITE) AND 5-FLUOROCYTOSINE AND 5-FLUOROSALICYLIC ACID HYDRATE (GREEN) AND RIGHT, THE TWO STRUCTURES SLOWLY MOVING OUT OF SYNC.....	192
FIGURE 4.36. THE TWO INDEPENDENT HETERODIMERS IN THE MOLECULAR COMPLEX OF CYTOSINIUM 3-HYDROXYBENZOATE HEMIHYDRATE MOLECULAR COMPLEX.....	193
FIGURE 4.37. TOP, THE TWO INDEPENDENT HYDROGEN BONDED HETERODIMERS SHOWING THE INCREASED TORSION ANGLE OF THE CARBOXYLIC ACID GROUP OF ONE (TOP RIGHT) COMPARED TO THE OTHER (TOP LEFT). BOTTOM, THE ASYMMETRIC UNIT OF THE CYTOSINIUM 3-HYDROXYBENZOATE MOLECULAR COMPLEX.....	193
FIGURE 4.38. THE TWO DISTINCT HYDROGEN BONDING ENVIRONMENTS OF THE INDEPENDENT CYTOSINIUM MOLECULES IN THE CYTOSINIUM 3-HYDROXYBENZOATE HEMIHYDRATE MOLECULAR COMPLEX.	194
FIGURE 4.39. ALTERNATING DIMER 1 (RED) AND 2 (GREEN) ALONG AN EXTENDED CHAIN IN THE CYTOSINIUM 3-HYDROXYBENZOATE HEMIHYDRATE MOLECULAR COMPLEX.	195
FIGURE 4.40. WATER MOLECULE ACTING AS A LINK BETWEEN THREE CYTOSINIUM AND 3-HYDROXYBENZOATE HETERODIMERS.....	195
FIGURE 4.41. WATER MOLECULE ACTING AS A LINK BETWEEN THREE CYTOSINIUM AND 3-HYDROXYBENZOATE HETERODIMERS.....	196
FIGURE 4.42. CHAINS FORMED WITH (A) JUST DIMER 1 (RED FOR CYTOSINIUM AND GREEN FOR THE 3-HYDROXYBENZOATE)AND (B) JUST DIMER 2 (YELLOW FOR CYTOSINIUM BLUE FOR THE 3-	

HYDROXYBENZOATE AND PURPLE FOR THE WATER MOLECULES) IN THE CYTOSINIUM AND 3-HYDROXYBENZOATE HEMIHYDRATE.	196
FIGURE 4.43. PACKING IN THE CYTOSINIUM 3-HYDROXYBENZOATE HEMIHYDRATE.	197
FIGURE 4.44. COMBINATION OF TYPE A AND B PSEUDO-BASE PAIRED HYDROGEN BONDING IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 3-HYDROXYBENZOIC ACID.....	199
FIGURE 4.45. LEFT. THE BASE-PAIRED 5-FLUOROCYTOSINE CHAIN; RIGHT, 3-HYDOXYBENZOIC ACID MOLECULES TIED TO EACH SIDE OF THE CHAIN IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 3-HYDROXYBENZOIC ACID.....	200
FIGURE 4.46. LEFT, CHANGE IN ORIENTATION OF ADJACENT CHAINS AND RIGHT, FLUORINE INTERACTIONS IN THE 5-FLUOROCYTOSINE AND 3-HYDROXYBENZOIC ACID MOLECULAR COMPLEX. GREEN REPRESENTS THE CHAIN HELD IN PLACE BY HYDROGEN BONDING AND RED REPRESENTS THE CHAIN HELD IN PLACE VIA FLUORINE.	200
FIGURE 4.47. STACKING INTERACTIONS IN THE 5-FLUOROCYTOSINE AND 3-HYDROXYBENZOIC ACID MOLECULAR COMPLEX.....	201
FIGURE 4.48 STAGGERED STACKING INTERACTIONS CORRESPONDING TO THE CLOSEST INTERACTION IN THE 5-FLUOROCYTOSINE AND 3-HYDROXYBENZOIC ACID MOLECULAR COMPLEX.....	201
FIGURE 4.49. CRISS-CROSS PATTERN FORMED IN THE 5-FLUOROCYTOSINE AND 3-HYDROXYBENZOIC ACID MOLECULAR COMPLEX.....	202
FIGURE 4.50. LEFT, THE ASYMMETRIC UNIT SHOWING PROTON TRANSFER FROM ONE 4-HYDROXYBENZOIC ACID MOLECULE TO ONE CYTOSINE MOLECULE AND RIGHT, THE PSEUDO-WATSON-CRICK BASE PAIR FOUND IN THE 2:3:2 MOLECULAR COMPLEX OF CYTOSINE 4-HYDROXYBENZOIC ACID HYDRATE.....	203
FIGURE 4.51. THE HYDROGEN BONDED CHAIN FORMED IN THE 4-HYDROXYBENZOIC ACID AND CYTOSINE HYDRATE MOLECULAR STRUCTURE.	204
FIGURE 4.52. EXTENDED CHAIN WITH WATER MOLECULES TIED TO ONE SIDE OF THE CHAIN IN THE CYTOSINE 4-HYDROXYBENZOIC ACID HYDRATE MOLECULAR COMPLEX.....	204
FIGURE 4.53. THE TWO FOUR MEMBERED HYDROGEN BONDED RINGS OF TWO WATER MOLECULES AND TWO 4-HYDROXYBENZOATE MOLECULES IN THE CYTOSINE 4-HYDROXYBENZOIC ACID HYDRATE MOLECULAR COMPLEX.....	205
FIGURE 4.54. CHARGED 3-HYDROXYBENZOIC ACID MOLECULE LINKING THE LAYERS IN THE CYTOSINE 4-HYDROXYBENZOIC ACID HYDRATE MOLECULAR COMPLEX.....	205
FIGURE 4.55. COMBINATION OF TWO FOUR MEMBERED RINGS FORMING AN EXTENDED CHAIN IN THE CYTOSINE AND 4-HYDROXYBENZOIC ACID HYDRATE MOLECULAR COMPLEX.....	206
FIGURE 4.56. STACKING INTERACTIONS IN THE CYTOSINE AND 4-HYDROXYBENZOIC ACID HYDRATE MOLECULAR COMPLEX.....	206
FIGURE 4.57. PACKING IN THE CYTOSINE AND 4-HYDROXYBENZOIC ACID HYDRATE MOLECULAR COMPLEX.....	206
FIGURE 4.58. LEFT, TYPE B HYDROGEN BONDING MOTIF AND RIGHT, FOUR MOLECULE BASIC BUILDING BLOCK IN THE 5-FLUOROCYTOSINE AND 4-HYDROXYBENZOIC ACID MOLECULAR COMPLEX.....	207
FIGURE 4.59. LEFT. TWO SEPARATE BUILDING BLOCKS LINKED VIA A SINGLE HYDROGEN BOND (BLACK LINES) AND RIGHT, ORIENTATION OF	

THE MOLECULES HYDROGEN BONDED TO FOUR MOLECULE BUILDING BLOCK VIA SINGLE HYDROGEN BONDS IN THE 5-FLUOROCYTOSINE AND 4-HYDROXYBENZOIC ACID MOLECULAR COMPLEX.....	208
FIGURE 4.60. LEFT. LAYERED NATURE OF STRUCTURE AND RIGHT, STACKING INTERACTIONS IN STRUCTURE OF 5-FLUOROCYTOSINE AND 4-HYDROXYBENZOIC ACID MOLECULAR COMPLEX.	208
FIGURE 4.61. LEFT, WEAK C-H...F HYDROGEN BOND AND RIGHT, WEAK C-H...O HYDROGEN BOND IN THE 5-FLUOROCYTOSINE AND 4-HYDROXYBENZOIC ACID MOLECULAR COMPLEX.....	209
FIGURE 4.62. STACKING INTERACTIONS IN THE 5-FLUOROCYTOSINE AND 4-HYDROXYBENZOIC ACID MOLECULAR COMPLEX. THREE STACKING INTERACTIONS SHOW WITH TWO BEING EQUIVALENT.	209
FIGURE 4.63. CHAINS OF 5-FLUOROCYTOSINE WITH ALTERNATING TYPE A AND B HYDROGEN BOND MOTIFS IN THE MOLECULAR COMPLEXES OF 5-FLUOROCYTOSINE WITH BENZOIC ACID (LEFT) AND 3-HYDROXYBENZOIC ACID (RIGHT).....	212
FIGURE 4.64. BENZOIC ACID (LEFT) AND 3-HYDROXYBENZOIC ACID (RIGHT) TIED TO THE 5-FLUOROCYTOSINE CHAINS IN THEIR RESPECTIVE STRUCTURES.	212
FIGURE 4.65. THE SIMILAR HYDROGEN BONDED UNITS INVOLVING C-H...F HYDROGEN BONDS IN THE LEFT, 5-FLUOROCYTOSINE BENZOIC ACID AND RIGHT, 5-FLUOROCYTOSINE 3-HYDROXYBENZOIC ACID MOLECULAR COMPLEXES.	213
FIGURE 4.66. TERMINATION OF 5-FLUOROCYTOSINE CHAIN VIA FORMATION OF A HYDROGEN BONDED RING IN THE 5-FLUOROCYTOSINE AND 4-HYDROXYBENZOIC ACID MOLECULAR COMPLEX.....	213
FIGURE 4.67. PRIMARY HYDROGEN BONDING MOTIF IN THE FOUR RELATED CHARGED STRUCTURES; (A) LEFT, CYTOSINIUM SALICYLATE MIDDLE, 5-FLUOROCYTOSINIUM SALICYLATE RIGHT, CYTOSINIUM 5-FLUOROSALICYLATE; (B) TWO INDEPENDENT HETERODIMERS IN CYTOSINIUM 3-HYDROXYBENZOATE HEMIHYDRATE.....	214
FIGURE 4.68. HYDROGEN BONDING ENVIRONMENT OF THE CYTOSINIUM MOLECULE IN: LEFT, CYTOSINIUM SALICYLATE; MIDDLE, 5-FLUOROCYTOSINIUM SALICYLATE; AND RIGHT, CYTOSINIUM 5-FLUOROSALICYLATE.	215
FIGURE 4.69. HYDROGEN BONDED CHAINS OF HETERODIMERS IN: LEFT, CYTOSINIUM SALICYLATE; MIDDLE; CYTOSINIUM 5-FLUOROSALICYLATE; RIGHT, 5-FLUOROCYTOSINIUM SALICYLATE.....	215
FIGURE 4.70. THE LINKING OF THE CYTOSINIUM CHAINS IN THE MOLECULAR COMPLEXES OF: TOP, LEFT CYTOSINIUM SALICYLATE; TOP RIGHT, CYTOSINIUM 5-FLUOROSALICYLATE; BOTTOM, 5-FLUOROCYTOSINIUM SALICYLATE.	216
FIGURE 4.71. FLUORINE INTERACTIONS IN LEFT, CYTOSINIUM 5-FLUOROSALICYLATE AND RIGHT 5-FLUOROCYTOSINIUM SALICYLATE MOLECULAR COMPLEXES.	216
FIGURE 4.72. THE HYDROGEN BONDS FROM THE HETERODIMER (DIMER 1 GREEN AND DIMER 2 BLUE) IN THE CYTOSINIUM 3-HYDROXYBENZOATE HEMIHYDRATE MOLECULAR COMPLEX	217
FIGURE 4.73. PSEUDO-WATSON-CRICK BASE PAIR IN THE, LEFT CYTOSINE 4-HYDROXYBENZOIC ACID HYDRATE MOLECULAR COMPLEX, MIDDLE 5-FLUOROCYTOSINE 5-FLUOROSALICYLIC ACID HYDRATE MOLECULAR COMPLEX AND RIGHT, CYTOSINE 5-FLUOROSALICYLIC ACID HYDRATE MOLECULAR COMPLEX.....	218

FIGURE 4.74 LEFT, HYDROGEN BONDED CHAINS AND RIGHT, LINKING OF THE CHAINS IN THE MOLECULAR COMPLEX OF CYTOSINE AND BENZOIC ACID.....	219
FIGURE 4.75. PSEUDO-WATSON-CRICK AND TYPE B BASE-PAIRING IN THE CYTOSINE AND 5-FLUOROSALICYLIC ACID MOLECULAR COMPLEX (LEFT) AND THE 5-FLUOROCYTOSINE AND 5-FLUOROSALICYLIC ACID MOLECULAR COMPLEX (RIGHT).	220
FIGURE 4.76 TERMINATION OF THE CHAIN IN THE CYTOSINE AND 5-FLUOROSALICYLIC ACID MOLECULAR COMPLEX (LEFT) AND 5-FLUOROCYTOSINE AND 5-FLUOROSALICYLIC ACID MOLECULAR COMPLEX (RIGHT).....	220
FIGURE 4.77. EXTENDED HYDROGEN BONDED CHAIN OF ALTERNATING CYTOSINE AND 4-HYDROXYBENZOIC ACID MOLECULES IN THE CYTOSINE AND 4-HYDROXYBENZOIC ACID MOLECULAR COMPLEX....	221
FIGURE 4.78. STAGGERED STACKING INTERACTIONS OBSERVED IN SEVERAL STRUCTURES IN THIS CHAPTER.....	222

Chapter 5

FIGURE 5.1. THE HETERO BASE PAIRING BETWEEN 5-FLUOROURACIL AND UREA IN THE 1:1 MOLECULAR COMPLEX.....	223
FIGURE 5.2. HETERO BASE- PAIRED CHAINS OF UREA AND 5-FLUOROURACIL.	224
FIGURE 5.3. 10 MEMBERED RING CREATED BY HYDROGEN BONDS IN THE MOLECULAR COMPLEX OF 5-FLUOROURACIL AND UREA.....	225
FIGURE 5.4 NET-LIKE STRUCTURE IN THE 5-FLUOROURACIL AND UREA MOLECULAR COMPLEX. THE BLACK BOX REPRESENTS THE CHAIN FORMED IN FIGURE 5.3 AND THE RED BOX REPRESENTS ONE OF THE TWO CHAINS IN FIGURE 5.3.	225
FIGURE 5.5. LEFT, FLUORINE=FLUORINE AND BASE-STACKING INTERACTIONS HOLDING THE INVERTED CHAIN IN THE CENTRE OF THE TEN MOLECULE HYDROGEN BONDED RING; RIGHT THE CRYSTAL PACKING VIEWED ALONG THE C AXIS IN THE 5-FLUOROURACIL AND UREA MOLECULAR COMPLEX.	225
FIGURE 5.6. THE INTERTWINED NETS IN THE MOLECULAR COMPLEX OF 5-FLUOROURACIL WITH UREA. THE FIRST NET IS FORMED BETWEEN THE GREEN AND ORANGE CHAINS, AND ITS INVERTED EQUIVALENT IS BETWEEN THE WHITE AND PURPLE.	226
FIGURE 5.7. FLUORINE INTERACTIONS IN THE 5-FLUOROURACIL AND UREA MOLECULAR COMPLEX.....	226
FIGURE 5.8. STACKING INTERACTIONS IN THE 5-FLUOROURACIL AND UREA MOLECULAR COMPLEX.....	227
FIGURE 5.9. GLOBAL PACKING IN THE 5-FLUOROURACIL AND UREA MOLECULAR COMPLEX. THE 5-FLUOROURACIL MOLECULES ARE SHOWN IN RED AND THE UREA MOLECULES IN GREEN.	227
FIGURE 5.10. THE HYDROGEN BONDING ENVIRONMENT OF THE URACIL MOLECULE IN ITS MOLECULAR COMPLEX WITH UREA.	228
FIGURE 5.11. HETERO BASE- PAIRED CHAINS OF UREA AND URACIL.	229
FIGURE 5.12. 10 MEMBERED RING CREATED BY HYDROGEN BONDS IN THE MOLECULAR COMPLEX OF URACIL AND UREA.	229
FIGURE 5.13. LEFT, THE NET LIKE STRUCTURE AND RIGHT, THE CRYSTAL PACKING OF INTERTWINED NETS VIEWED ALONG THE C AXIS IN THE URACIL AND UREA MOLECULAR COMPLEX.	230

FIGURE 5.14. THE INTERTWINED NETS IN THE MOLECULAR COMPLEX OF URACIL WITH UREA. THE FIRST NET IS FORMED BETWEEN THE GREEN AND ORANGE CHAINS, AND ITS INVERTED EQUIVALENT IS BETWEEN THE WHITE AND PURPLE.	230
FIGURE 5.15. STACKING INTERACTIONS IN THE URACIL AND UREA MOLECULAR COMPLEX.....	231
FIGURE 5.16. GLOBAL PACKING IN THE URACIL AND UREA MOLECULAR COMPLEX. THE URACIL MOLECULES ARE SHOWN IN RED AND THE UREA MOLECULES IN GREEN.	231
FIGURE 5.17. HOMO BASE-PAIRING IN THE 5-FLUOROURACIL AND THIOUREA MOLECULAR COMPLEX.....	232
FIGURE 5.18. ALTERNATING PAIRS OF 5-FLUOROURACIL AND THIOUREA MOLECULES IN THE HYDROGEN BONDED CHAIN.....	232
FIGURE 5.19. THE HYDROGEN BONDING PATTERN IN THE CHAINS OF ALTERNATING 5-FLUOROURACIL AND THIOUREA MOLECULES.	233
FIGURE 5.20. THE TWO TYPES OF BASE-PAIRING IN THE MOLECULAR COMPLEX OF 5-FLUOROURACIL WITH THIOUREA. BLACK INDICATES THE HOMO BASE PAIR AND RED REPRESENTS THE THIOUREA 5-FLUOROURACIL HETERO BASE PAIR MOTIF.	233
FIGURE 5.21. WEAK HYDROGEN BOND INVOLVING THE FLUORINE ATOM IN THE 5-FLUOROURACIL AND THIOUREA MOLECULAR COMPLEX.....	234
FIGURE 5.22. WEAK HYDROGEN BOND LINKING CHAINS OF 5-FLUOROURACIL AND THIOUREA.....	234
FIGURE 5.23. SULPHUR INTERACTIONS BETWEEN THE LAYERS IN THE MOLECULAR COMPLEX OF 5-FLUOROURACIL AND THIOUREA.	235
FIGURE 5.24 OVERLAYED CHAINS IN THE MOLECULAR COMPLEX OF 5-FLUOROURACIL AND THIOUREA.....	235
FIGURE 5.25. EXTENDED 5-FLUOROCYTOSINE CHAIN (CHAIN 1) FORMED IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND THIOUREA HYDRATE.....	236
FIGURE 5.26. THE TWO INDEPENDENT 5-FLUOROCYTOSINE MOLECULES (ONE GREEN AND ONE RED) COMBINE IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND THIOUREA HYDRATE TO FORM AN EXTENDED CHAIN (CHAIN 2).	237
FIGURE 5.27. ALL HYDROGEN BONDS (LABELLED) IN THE 5-FLUOROCYTOSINE CHAIN 2 IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND THIOUREA HYDRATE.....	237
FIGURE 5.28. THIOUREA LINKED TO EITHER SIDE OF THE 5-FLUOROCYTOSINE CHAIN (CHAIN 2) IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND THIOUREA HYDRATE.	238
FIGURE 5.29. THIOUREA LINKING SEPARATE LAYERS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND THIOUREA HYDRATE.	238
FIGURE 5.30. CHAINS OVERLAYED IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND THIOUREA HYDRATE.....	239
FIGURE 5.31. HYDROGEN BONDED RING FORMED BETWEEN TWO THIOUREA MOLECULES AND TWO WATER MOLECULES IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND THIOUREA HYDRATE.	239
FIGURE 5.32. LEFT, HYDROGEN BONDED RING FORMED BETWEEN TWO THIOUREA MOLECULES. RIGHT, THE HYDROGEN BONDED RING BETWEEN TWO THIOUREA MOLECULES LINKING FOUR MEMBERED RINGS TOGETHER IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND THIOUREA HYDRATE.....	240

FIGURE 5.33. WATER MOLECULES TIED TO EITHER SIDE OF THE 5-FLUOROCYTOSINE CHAIN (CHAIN 1) IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND THIOUREA HYDRATE.	240
FIGURE 5.34. WATER MOLECULES (YELLOW) LINKING THE SEPARATE LAYERS TOGETHER AND THE TWO INDEPENDENT CHAINS IN THE STRUCTURE (CHAIN 1 PURPLE AND CHAIN 2 GREEN) IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND THIOUREA HYDRATE.....	241
FIGURE 5.35. LEFT, FLUORINE INTERACTIONS BETWEEN LAYERS AND RIGHT, THE FLUORINE INTERACTIONS (RED BOX) AND THE FOUR MEMBERED RING OF THIOUREA AND WATER LINKING SEPARATE LAYERS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND THIOUREA HYDRATE.	242
FIGURE 5.36. THE BASE PAIRED CHAINS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND UREA.	242
FIGURE 5.37. UREA HYDROGEN BONDED CHAINS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND UREA.....	243
FIGURE 5.38. FLUORINE INTERACTIONS LINKING SEPARATE CHAINS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND UREA.....	244
FIGURE 5.39. COLOUR CODED STRUCTURE WITH THE CYTOSINE MOLECULES LABELLED BLUE AND GREEN AND THE UREA MOLECULES LABELLED RED SHOWING THE POSITIONING OF THE UREA CHAINS IN RELATION TO THE 5-FLUOROCYTOSINE CHAINS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND UREA	244
FIGURE 5.40. HYDROGEN BONDED CHAINS IN LEFT, THE 5-FLUOROURACIL AND UREA MOLECULAR COMPLEX AND RIGHT, THE URACIL AND UREA MOLECULAR COMPLEX.....	246
FIGURE 5.41. HYDROGEN BONDING BETWEEN CHAINS IN THE MOLECULAR COMPLEXES OF LEFT, 5-FLUOROURACIL AND UREA AND RIGHT, URACIL AND UREA.	247
FIGURE 5.42. FLUORINE INTERACTION (RED DOTTED LINE) AND BIFURCATED HYDROGEN BOND (BLUE DOTTED LINE) IN THE MOLECULAR COMPLEX OF 5-FLUOROURACIL AND UREA.....	247
FIGURE 5.43. STRUCTURE OVERLAY OF THE MOLECULAR COMPLEXES OF 5-FLUOROURACIL AND UREA (GREEN) AND URACIL AND UREA (WHITE) SHOWING DIVERGENCE OVER LONGER LENGTH SCALES.	248
FIGURE 5.44. THE INTERACTIONS INVOLVING THE FLUORINE ATOM IN THE MOLECULAR COMPLEX OF 5-FLUOROURACIL AND UREA.....	248
FIGURE 5.45. OVERLAYED STRUCTURES OF URACIL AND UREA (RED) AND THE 5-FLUOROURACIL AND UREA (GREEN) SHOWING SUBTLE INTERACTIONS DUE TO THE FLUORINE ATOM.....	248
FIGURE 5.46. LEFT HETERO- PSEUDO-BASE PAIRING IN THE 5-FLUOROURACIL AND UREA MOLECULAR COMPLEX AND RIGHT, HOMO-BASE PAIRING IN THE 5-FLUOROURACIL AND THIOUREA MOLECULAR COMPLEX.....	249
FIGURE 5.47. LEFT, EXTENDED CHAIN IN THE 5-FLUOROURACIL AND UREA MOLECULAR COMPLEX AND RIGHT, EXTENDED CHAIN IN THE 5-FLUOROURACIL AND THIOUREA MOLECULAR COMPLEX.....	250
FIGURE 5.48. HETERO-BASE PAIR HYDROGEN BONDING MOTIFS, LEFT ON BOTH SIDES OF 5-FLUOROURACIL MOLECULE IN THE 5-FLUOROURACIL AND UREA MOLECULAR COMPLEX AND RIGHT, ONLY ON ONE SIDE OF THE 5-FLUOROURACIL MOLECULE IN THE 5-FLUOROURACIL AND THIOUREA MOLECULAR COMPLEX.	250

FIGURE 5.48. LEFT, THIOUREA ACTING AS A LINKER IN THE 5-FLUOROCYTOSINE AND THIOUREA HYDRATE MOLECULAR COMPLEX AND RIGHT UREA CHAIN BUILT UP IN THE 5-FLUOROCYTOSINE AND UREA MOLECULAR COMPLEX.	251
FIGURE 5.48. FLUORINE HYDROGEN BONDS IN THE MOLECULAR COMPLEXES OF LEFT. 5-FLUOROCYTOSINE AND THIOUREA HYDRATE AND RIGHT, 5-FLUOROCYTOSINE AND UREA.	252

Chapter 6

FIGURE 6.1. HETERODIMER FORMED THROUGH A HYDROGEN-BONDED RING BETWEEN 5-FLUOROCYTOSINIUM AND 2,6-DIHYDROXYBENZOATE MOLECULES IN FORM I.	254
FIGURE 6.2. HYDROGEN BONDED HETERODIMERS ASSEMBLING INTO CORRUGATED LAYERS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINIUM 2,6-DIHYDROXYBENZOATE FORM I.	255
FIGURE 6.3. BASE-PAIR STYLE MOTIF INVOLVING FLUORINE IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINIUM 2,6-DIHYDROXYBENZOATE FORM I.	255
FIGURE 6.4. FLUORINE INTERACTIONS LINKING BASE PAIRS IN PARALLEL STEPPED LAYERS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINIUM 2,6-DIHYDROXYBENZOATE FORM I.	255
FIGURE 6.5. STACKING INTERACTIONS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINIUM 2,6-DIHYDROXYBENZOATE FORM I (5-FLUOROCYTOSINIUM IN GREEN AND 2,6-DIHYDROXYBENZOATE IN BLUE).	256
FIGURE 6.6. HYDROGEN-BONDED RING FORMED BETWEEN 5-FLUOROCYTOSINIUM AND 2,6-DIHYDROXYBENZOATE IN THE FORM II MOLECULAR COMPLEX.	256
FIGURE 6.7. AN OVERLAY OF THE HETERODIMERS OF THE FORM I (GREEN) AND FORM II (WHITE) MOLECULAR COMPLEXES OF 5-FLUOROCYTOSINIUM 2,6-DIHYDROXYBENZOATE.	257
FIGURE 6.8. HYDROGEN BONDED HETERODIMERS ASSEMBLING INTO PLANAR LAYERS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINIUM 2,6-DIHYDROXYBENZOATE FORM II.	257
FIGURE 6.9. FLUORINE INTERACTIONS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINIUM 2,6-DIHYDROXYBENZOATE FORM II.	258
FIGURE 6.10. STACKING INTERACTIONS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINIUM 2,6-DIHYDROXYBENZOATE FORM II.	258
FIGURE 6.11. AN OVERLAY OF THE MOLECULAR COMPLEXES OF FORM I (GREEN) AND FORM II (WHITE) OF 5-FLUOROCYTOSINIUM 2,6-DIHYDROXYBENZOATE SHOWING THE DIFFERENT CONSTRUCTION OF THE LAYERS.	259
FIGURE 6.12. CHAINS OF ALTERNATING TYPE B AND PSEUDO-WATSON-CRICK BASE PAIRS IN THE HYDRATED MOLECULAR COMPLEX OF CYTOSINE AND 2,6-DIHYDROXYBENZOIC ACID.	260
FIGURE 6.13. THE HYDROGEN BONDING TO THE CYTOSINE CHAINS IN THE MOLECULAR COMPLEX OF CYTOSINE AND 2,6-DIHYDROXYBENZOIC ACID HYDRATE. THE BENZOIC ACID MOLECULES LIE ON ONE SIDE OF THE CHAIN AND THE WATER MOLECULES ON THE OTHER SIDE.	261
FIGURE 6.14. HYDROGEN BONDING INVOLVING THE WATER MOLECULE IN THE CYTOSINE AND 2,6-DIHYDROXYBENZOIC ACID HYDRATE MOLECULAR COMPLEX.	261

FIGURE 6.15. HERRING-BONE ARRANGEMENT ALONG THE A-AXIS IN THE MOLECULAR COMPLEX OF CYTOSINE AND 2,6-DIHYDROXYBENZOIC ACID HYDRATE.	262
FIGURE 6.16. STACKING INTERACTIONS IN THE MOLECULAR COMPLEX OF CYTOSINE AND 2,6-DIHYDROXYBENZOIC ACID HYDRATE.	262
FIGURE 6.17. TYPE B BASE-PAIRING IN THE CYTOSINIUM 2,4-DIHYDROXYBENZOATE HYDRATE FORM I MOLECULAR COMPLEX.....	263
FIGURE 6.18. TERMINATION OF CYTOSINE CHAINS VIA HYDROGEN-BONDED RING FORMATION CREATING A FOUR MOLECULE UNIT IN THE CYTOSINIUM 2,4-DIHYDROXYBENZOATE HYDRATE FORM I MOLECULAR COMPLEX.....	264
FIGURE 6.19. LEFT, THE THREE HYDROGEN BONDS INVOLVING WATER MOLECULES AND RIGHT. THE WATER MOLECULE CONNECTING THE LAYERS IN THE CYTOSINIUM 2,4-DIHYDROXYBENZOATE HYDRATE FORM I MOLECULAR COMPLEX.	265
FIGURE 6.20. LINK BETWEEN FOUR MOLECULE BLOCKS IN THE CYTOSINIUM 2,4-DIHYDROXYBENZOATE HYDRATE MOLECULAR COMPLEX.....	265
FIGURE 6.21. THE LAYERED STRUCTURE OF THE CYTOSINIUM 2,4-DIHYDROXYBENZOATE HYDRATE FORM I MOLECULAR COMPLEX.....	265
FIGURE 6.22. STACKING INTERACTIONS IN THE MOLECULAR COMPLEX OF CYTOSINIUM 2,4-DIHYDROXYBENZOATE HYDRATE FORM I. CYTOSINIUM (GREEN), 2,4-DIHYDROXYBENZOATE (BLUE) AND WATER (RED).	266
FIGURE 6.23. TYPE B BASE-PAIRING IN THE CYTOSINE 2,4-DIHYDROXYBENZOIC ACID HYDRATE FORM II MOLECULAR COMPLEX.	267
FIGURE 6.24. TERMINATION OF THE CYTOSINIUM CHAIN VIA HYDROGEN-BONDED RING FORMATION WITH THE 2,4-DIHYDROXYBENZOATE MOLECULES IN THE CYTOSINIUM 2,4-DIHYDROXYBENZOATE HYDRATE FORM II MOLECULAR COMPLEX.....	267
FIGURE 6.25. HYDROGEN BONDS INVOLVING WATER MOLECULES IN THE CYTOSINIUM 2,4-DIHYDROXYBENZOATE HYDRATE FORM II MOLECULAR COMPLEX.....	267
FIGURE 6.26. DIFFERENCES IN MOLECULES BONDED TO THE FOUR MOLECULE UNIT (RED) IN THE MOLECULAR COMPLEXES OF CYTOSINIUM 2,4-DIHYDROXYBENZOATE HYDRATE FORM I (LEFT) AND FORM II (RIGHT).....	269
FIGURE 6.27. STACKING INTERACTIONS IN THE MOLECULAR COMPLEX OF CYTOSINIUM 2,4-DIHYDROXYBENZOATE HYDRATE FORM II.....	269
FIGURE 6.28. LEFT. FOUR-MOLECULE HYDROGEN BONDED UNIT AND RIGHT, TWO FOUR MOLECULE HYDROGEN BONDED RINGS LINKED BY TYPE B BASE PAIRING IN THE 5-FLUOROCYTOSINIUM 3,5-DIHYDROXYBENZOATE HYDRATE MOLECULAR COMPLEX.	270
FIGURE 6.29. FOUR MOLECULE HYDROGEN BONDED RING WITH BRIDGING WATER MOLECULE IN THE 5-FLUOROCYTOSINIUM 3,5-DIHYDROXYBENZOATE HYDRATE MOLECULAR COMPLEX.	271
FIGURE 6.30. WATER MOLECULE LINKING LAYERS IN THE 5-FLUOROCYTOSINIUM 3,5-DIHYDROXYBENZOATE HYDRATE MOLECULAR COMPLEX.....	271
FIGURE 6.31. RINGS FORMED WITH WATER MOLECULES AND 3,5-DIHYDROXYBENZOATE MOLECULES IN THE 5-FLUOROCYTOSINIUM 3,5-DIHYDROXYBENZOATE HYDRATE MOLECULAR COMPLEX.	272

FIGURE 6.32. THE CHANGES IN THE ORIENTATION OF THE WATER MOLECULES HELPING TO LINK THE IN THE 5-FLUOROCYTOSINIUM 3,5-DIHYDROXYBENZOATE HYDRATE MOLECULAR COMPLEX. GREEN AND PURPLE WATER MOLECULES IN SAME PLANE AS THE CHAINS AND THE BLACK, BLUE AND ORANGE MOLECULES SAT BETWEEN PLANES CONNECTING SEPARATE CHAINS.....	273
FIGURE 6.33. STACKING INTERACTIONS IN THE 5-FLUOROCYTOSINIUM 3,5-DIHYDROXYBENZOATE MOLECULAR COMPLEX.	273
FIGURE 6.34. TWO INDEPENDENT HYDROGEN BONDED RINGS FORMED IN THE MOLECULAR COMPLEX OF CYTOSINIUM 3,5-DIHYDROXYBENZOATE HEMIHYDRATE. LEFT, THE 3,5-DIHYDROXYBENZOATE ADOPTS A SYMMETRIC CONFIGURATION, AND RIGHT, THE 3,5-DIHYDROXYBENZOATE ADOPTS AN ASYMMETRIC CONFIGURATION.	274
FIGURE 6.35. TWO DIFFERENT 3,5-DIHYDROXYBENZOATE MOLECULES IN THE CYTOSINIUM 3,5-DIHYDROXYBENZOATE MOLECULAR COMPLEX.	275
FIGURE 6.36. LINKING OF TWO HETERODIMERS THROUGH AN N-H...O HYDROGEN BOND FROM A CYTOSINIUM TO A 3,5-DIHYDROXYBENZOATE MOLECULE.....	275
FIGURE 6.37. WATER MOLECULE LINKING TWO SYMMETRIC 3,5-DIHYDROXYBENZOATE MOLECULES TO AN ASYMMETRIC DIMER IN THE CYTOSINIUM 3,5-DIHYDROXYBENZOATE HEMIHYDRATE MOLECULAR COMPLEX.	276
FIGURE 6.38. HYDROGEN BONDS INVOLVING THE ASYMMETRIC DIMER IN THE CYTOSINIUM 3,5-DIHYDROXYBENZOATE HEMIHYDRATE MOLECULAR COMPLEX.....	276
FIGURE 6.39. THE DIFFERENT ORIENTATIONS OF THE TWO HETERODIMERS IN THE CYTOSINIUM 3,5-DIHYDROXYBENZOATE HEMIHYDRATE MOLECULAR COMPLEX.....	277
FIGURE 6.40. STACKING INTERACTIONS IN THE CYTOSINIUM 3,5-DIHYDROXYBENZOATE HEMIHYDRATE MOLECULAR COMPLEX. RED AND BLUE ARE THE 3,5-DIHYDROXYBENZOATE AND CYTOSINIUM MOLECULES IN A SYMMETRIC DIMER, RESPECTIVELY AND GREEN AND ORANGE ARE THE 3,5-DIHYDROXYBENZOATE AND CYTOSINIUM MOLECULES IN AN ASYMMETRIC DIMER, RESPECTIVELY. WATER IS YELLOW.....	278
FIGURE 6.41. HYDROGEN BONDS INVOLVING THE 5-FLUOROURACIL MOLECULE IN THE MOLECULAR COMPLEX OF 5-FLUOROURACIL 3,5-DIHYDROXYBENZOIC ACID HYDRATE.	279
FIGURE 6.42. WATER MOLECULES DISRUPTING THE TYPE B BASE-PAIRING IN THE MOLECULAR COMPLEX OF 5-FLUOROURACIL 3,5-DIHYDROXYBENZOIC ACID HYDRATE.	280
FIGURE 6.43. 5-FLUOROURACIL SURROUNDED BY A RING OF 3,5-DIHYDROXYBENZOIC ACID AND WATER MOLECULES IN MOLECULAR COMPLEX OF 5-FLUOROURACIL AND 3,5-DIHYDROXYBENZOIC ACID HYDRATE.....	280
FIGURE 6.44 RING 1 (LEFT) AND RING 2 (RIGHT) MADE UP OF 3,5-DIHYDROXYBENZOIC ACID AND WATER MOLECULES IN THE MOLECULAR COMPLEX OF 5-FLUOROURACIL AND 3,5-DIHYDROXYBENZOIC ACID HYDRATE.	281
FIGURE 6.45. RING 1 WITH LINK TO THE BASE PAIR (LEFT) AND RING 2 WITH LINK TO THE BASE PAIR (RIGHT) VIA WATER MOLECULES IN	

MOLECULAR COMPLEX OF 5-FLUOROURACIL AND 3,5-DIHYDROXYBENZOIC ACID HYDRATE.	282
FIGURE 6.46. HYDROGEN BONDS LINKING SEPARATE RINGS IN THE SAME PLANE IN THE MOLECULAR COMPLEX OF 5-FLUOROURACIL AND 3,5-DIHYDROXYBENZOIC ACID HYDRATE.	282
FIGURE 6.47. STACKING INTERACTIONS IN THE MOLECULAR COMPLEX OF 5-FLUOROURACIL AND 3,5-DIHYDROXYBENZOIC ACID HYDRATE.	283
FIGURE 6.48 PSEUDO-WATSON-CRICK UNIT (LEFT) AND THE 3,5-DINITROBENZOIC ACID DIMER (RIGHT) IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 3,5-DINITROBENZOIC ACID.	284
FIGURE 6.49. FOUR MOLECULE BUILDING BLOCK IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 3,5-DINITROBENZOIC ACID	285
FIGURE 6.50. 3,5-DINITROBENZOIC ACID DIMER LINKING SEPARATE FOUR MOLECULE BUILDING BLOCKS TOGETHER IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 3,5-DINITROBENZOIC ACID.	285
FIGURE 6.51. FLUORINE INTERACTIONS (LEFT) AND THE HERRINGBONE STRUCTURE PRODUCED (RIGHT) IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 3,5-DINITROBENZOIC ACID.	286
FIGURE 6.52 STACKING INTERACTIONS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 3,5-DINITROBENZOIC ACID.	286
FIGURE 6.53. SIMILAR BUILDING BLOCKS FOUND IN THE MOLECULAR COMPLEXES OF CYTOSINIUM 2,4-DIHYDROXYBENZOATE HYDRATE FORM I (LEFT), FORM II (MIDDLE) AND 5-FLUOROCYTOSINIUM 3,5-DIHYDROXYBENZOATE HYDRATE (RIGHT).	289
FIGURE 6.54. STEPS IN THE FOUR MOLECULE BUILDING BLOCKS FOUND IN THE MOLECULAR COMPLEXES OF CYTOSINE AND 2,4-DIHYDROXYBENZOIC ACID HYDRATE FORM I (LEFT), FORM II (MIDDLE), AND 5-FLUOROCYTOSINE AND 3,5-DIHYDROXYBENZOIC ACID HYDRATE (RIGHT).	289

Chapter 7

FIGURE 7.1 PSEUDO-WATSON-CRICK BASE PAIR MOTIF IN THE MOLECULAR COMPLEX OF CYTOSINE ISONICOTINIC ACID HYDRATE.....	292
FIGURE 7.2 BASE-PAIRED CHAIN OF CYTOSINE MOLECULES WITH ALTERNATING PSEUDO-WATSON-CRICK AND FOUR MOLECULE HYDROGEN BONDED RINGS IN THE MOLECULAR COMPLEX OF CYTOSINE ISONICOTINIC ACID HYDRATE	293
FIGURE 7.3 WATER AND ISONICOTINIC ACID MOLECULES TIED TO THE CYTOSINE CHAIN VIA HYDROGEN BONDS IN THE MOLECULAR COMPLEX OF CYTOSINE ISONICOTINIC ACID HYDRATE.....	294
FIGURE 7.4 WATER MOLECULE LINKING TWO ISONICOTINIC ACID MOLECULES VIA HYDROGEN BONDING IN THE MOLECULAR COMPLEX OF CYTOSINE ISONICOTINIC ACID HYDRATE.....	294
FIGURE 7.5 LEFT, LAYERS CONNECTED THROUGH HYDROGEN BONDING INVOLVING THE WATER AND RIGHT, THE STAGGERED STACKING OF CYTOSINE CHAINS IN THE MOLECULAR COMPLEX OF CYTOSINE ISONICOTINIC ACID HYDRATE.....	295
FIGURE 7.6 STACKING OF THE ISONICOTINIC ACID MOLECULES (YELLOW AND PURPLE) IN THE MOLECULAR COMPLEX OF CYTOSINE ISONICOTINIC ACID HYDRATE.....	295
FIGURE 7.7 COLOUR CODING OF THE FOUR INDEPENDENT MOLECULES IN THE MOLECULAR COMPLEX OF CYTOSINE ISONICOTINIC ACID	

HYDRATE TO SHOW THE OVERALL PACKING. ONE WATER MOLECULE (BLUE) ANOTHER WATER MOLECULE (BLACK), NICOTINIC ACID MOLECULE (PURPLE), NEUTRAL CYTOSINE (RED) AND CHARGED CYTOSINE (GREEN)	296
FIGURE 7.8 HETERODIMER WITH A HYDROGEN-BONDED RING MOTIF IN THE MOLECULAR COMPLEX OF CYTOSINIUM NICOTINATE MONOHYDRATE	297
FIGURE 7.9 LEFT, HYDROGEN BONDS THE HETERODIMER AND RIGHT, CHAINS OF HETERODIMERS LINKED BY SINGLE HYDROGEN BONDS IN THE CYTOSINIUM NICOTINATE MONOHYDRATE MOLECULAR COMPLEX.....	297
FIGURE 7.10 THE CONNECTION OF NEIGHBOURING HETERODIMER CHAINS IN THE MOLECULAR COMPLEX OF CYTOSINIUM NICOTINATE MONOHYDRATE. THE BIFURCATED HYDROGEN BOND INVOLVING THE WATER MOLECULE IS REPRESENTED BY GREEN DOTTED LINES AND THE WEAK C-H...O HYDROGEN BOND BY A VIOLET DOTTED LINE.	298
FIGURE 7.11 WATER MOLECULE CONNECTING THE LAYERS IN THE MOLECULAR COMPLEX OF CYTOSINIUM NICOTINATE MONOHYDRATE	298
FIGURE 7.12. EIGHT MOLECULE HYDROGEN BONDED RING IN THE MOLECULAR COMPLEX OF CYTOSINIUM NICOTINATE MONOHYDRATE	299
FIGURE 7.13 (A) STACKING INTERACTIONS AND (B) STAGGERED STACKING IN THE MOLECULAR COMPLEX OF CYTOSINIUM NICOTINATE MONOHYDRATE	299
FIGURE 7.14 COLOUR CODED STRUCTURE TO SHOW PLANAR NATURE AND THE ROLE OF THE WATER MOLECULE CONNECTING SEPARATE LAYERS IN THE MOLECULAR COMPLEX OF CYTOSINIUM NICOTINATE MONOHYDRATE. WATER (RED), CYTOSINE (GREEN) AND NICOTINIC ACID (YELLOW)	299
FIGURE 7.15. LEFT, PSEUDO-WATSON-CRICK BASE PAIR UNIT AND RIGHT, CHAINS OF CYTOSINE MOLECULES IN THE MOLECULAR COMPLEX OF CYTOSINE AND 2-NITROBENZOIC ACID.....	300
FIGURE 7.16 2-NITROBENZOATE MOLECULES HYDROGEN BONDED TO CYTOSINE BASE-PAIRED CHAIN IN THE 2:1 CYTOSINE AND 2-NITROBENZOIC ACID MOLECULAR COMPLEX.	301
FIGURE 7.17 2-NITROBENZOATE MOLECULES LINKING PARALLEL CYTOSINE BASE-PAIRED CHAINS VIA SINGLE HYDROGEN BONDS IN THE 2:1 CYTOSINE AND 2-NITROBENZOIC ACID MOLECULAR COMPLEX.....	302
FIGURE 7.18 LINKING BETWEEN CHAINS USING WEAK HYDROGEN BONDS AND THE SUBSEQUENT ORIENTATION OF THE CHAINS INVOLVED IN THE STRUCTURE OF CYTOSINE AND 2-NITROBENZOIC ACID	302
FIGURE 7.19 THE ORIENTATION OF THE BUILDING BLOCKS DESCRIBED SHOWING NON- PLANAR PACKING IN THE STRUCTURE OF CYTOSINE AND 2-NITROBENZOIC ACID	303
FIGURE 7.20 GLOBAL PACKING OF THE CYTOSINE AND 2-NITROBENZOIC ACID MOLECULAR COMPLEX. CYTOSINE MOLECULES ARE SHOWN IN RED AND 2-NITROBENZOATE MOLECULES IN GREEN.....	304
FIGURE 7.21.LEFT, PSEUDO-WATSON-CRICK BASE PAIRED HYDROGEN BONDING MOTIF AND RIGHT, CYTOSINE CHAINS IN THE MOLECULAR COMPLEX OF CYTOSINE AND 3-BROMOBENZOIC ACID.....	304

FIGURE 7.22 3-BROMOBENZOATE MOLECULES LINKING PARALLEL CYTOSINE CHAINS VIA SINGLE HYDROGEN BONDS IN THE MOLECULAR COMPLEX OF CYTOSINE AND 3-BROMOBENZOIC ACID.....	305
FIGURE 7.23 COMPARISON OF STRUCTURES OF (A) 3-BROMOBENZOIC ACID AND CYTOSINE AND (B) CYTOSINE AND 2-NITROBENZOIC ACID WITH CYTOSINE IN EACH STRUCTURE BEING REPRESENTED BY RED AND BENZOIC ACIDS REPRESENTED BY GREEN.....	306
FIGURE 7.24 HYDROGEN-BONDED RING MOTIF IN THE CYTOSINIUM 2,4,6-TRIMETHYLBENZOATE MONOHYDRATE MOLECULAR COMPLEX.	307
FIGURE 7.25 LEFT, CHAIN OF CYTOSINIUM MOLECULES LINKED BY SINGLE HYDROGEN BONDS AND RIGHT, HYDROGEN BONDS INVOLVING THE WATER MOLECULES IN THE MOLECULAR COMPLEX OF CYTOSINIUM 2,4,6-TRIMETHYLBENZOATE MONOHYDRATE.....	308
FIGURE 7.26 ORIENTATION (APPROXIMATELY PERPENDICULAR) OF NEIGHBOURING CYTOSINIUM CHAINS IN THE MOLECULAR COMPLEX OF CYTOSINIUM 2,4,6-TRIMETHYLBENZOATE MONOHYDRATE.....	308
FIGURE 7.27 STACKING OF CYTOSINIUM ON TOP OF 2,4,6-TRIMETHYLBENZOATE IN THE MOLECULAR COMPLEX OF CYTOSINIUM 2,4,6-TRIMETHYLBENZOATE MONOHYDRATE.....	309

Chapter 8

FIGURE 8.1. MOLECULAR STRUCTURES OF THE PURINE PAIR, ADENINE AND GUANINE, AND THE PYRIMIDINE PAIR, THYMINE AND CYTOSINE.....	313
FIGURE 8.2. PROPELLER TWISTING IN A DNA BASE PAIR [38].....	314
FIGURE 8.3. A SKETCH TO DEMONSTRATE BUCKLE IN A BASE PAIR MOTIF [39].....	314
FIGURE 8.4. THE FOUR TYPES OF BASE PAIRING FOR THE ADENINE AND THYMINE COMBINATION. (A) REVERSE HOOGSTEN (B) HOOGSTEN (C) REVERSE WATSON-CRICK (D) WATSON CRICK.	315
FIGURE 8.5. THE GROUP OF ATOMS ON THE THYMINE MOLECULE ARE CIRCLED IN RED AND THE GROUP OF ATOMS ON THE ADENINE MOLECULES ARE CIRCLED IN BLUE. CLOSE CONTACTS WERE DEFINED AS THE DISTANCE BETWEEN NEAREST ATOMS.....	316
FIGURE 8.6. DENDROGRAM OBTAINED FROM THE DSNAP PROGRAM FOR THE ADENINE AND THYMINE CSD SEARCH.	316
FIGURE 8.7. MMDS PLOT OF THE CLUSTERING OF ADENINE AND THYMINE BASE PAIRS OBTAINED FROM THE DSNAP PROGRAM.....	317
FIGURE 8.8. REVERSE HOOGSTEN HYDROGEN BONDING MOTIF IN THE RED AND YELLOW CLUSTERS [ADRBFT10] [142].	318
FIGURE 8.9. THE REVERSE HOOGSTEN BASE PAIR MOTIFS OF THE HITS IN THE GREEN CLUSTER. KECYUN (LEFT) AND MTYDAP (RIGHT).	319
FIGURE 8.10. PROPELLER TWIST IN THE MTYDAP [144] HIT OF THE GREEN CLUSTER OF THE ADENINE THYMINE DSNAP ANALYSIS.....	320
FIGURE 8.11. PRIMARY HYDROGEN BONDING MOTIF EXHIBITED BY ALL MEMBERS OF THE CYAN CLUSTER ILLUSTRATED BY ETABFU [148].	320
FIGURE 8.12. BARAAD [153] APPEARANCE IN THE BLUE AND GREEN STRIPED CLUSTERS. LEFT, UNIT IN THE BLUE CLUSTER SHOWING STACKING INTERACTIONS AND RIGHT, UNIT IN THE GREEN STRIPED CLUSTER SHOWING THE REVERSE WATSON CRICK MOTIF.	321
FIGURE 8.13. PRIMARY HYDROGEN BONDING MOTIF EXHIBITED BY ALL MEMBERS OF THE GREEN STRIPED CLUSTER ILLUSTRATED BY CETQUO [156].	322

FIGURE 8.14. SINGLE HYDROGEN BOND BETWEEN ADENINE AND THYMINE IN THE PINK CLUSTER ILLUSTRATED BY ETABFU [148].	323
FIGURE 8.15. PRIMARY BONDING MOTIF EXHIBITED BY ALL MEMBERS OF THE ORANGE STRIPED CLUSTER ILLUSTRATED BY EBPMUR [155].	323
FIGURE 8.16. THE SINGLE HYDROGEN BOND IN THE CYAN STRIPED (KECYUN [143]) THAT EXTENDS BASE PAIRED UNITS INTO A CHAIN.	325
FIGURE 8.17. BASE STACKING INTERACTIONS BETWEEN THE DEFINED AREAS OF THE STARTING MOLECULES IN THE BLUE STRIPED CLUSTER (FUJDET [157]).	326
FIGURE 8.18. BASE STACKING INTERACTIONS BETWEEN THE DEFINED AREAS OF THE STARTING MOLECULES IN THE PURPLE STRIPED CLUSTER (MIUDAP10 [151]).	326
FIGURE 8.19. DEFINED CONTACTS USED FOR SEARCH IN THE CSD BETWEEN CYTOSINE AND GUANINE MOLECULES.	327
FIGURE 8.20. DENDROGRAM OBTAINED FROM THE DSNAP PROGRAM WHEN WITH THE AUTOMATIC CUT LEVEL FOR CYTOSINE AND GUANINE FRAGMENTS.	328
FIGURE 8.21. MMDS PLOT OF RESULTS OBTAINED FROM THE DSNAP PROGRAM USING THE CALCULATED CUT LEVEL FOR THE CYTOSINE GUANINE HYDROGEN BONDED FRAGMENTS.	328
FIGURE 8.22. DENDROGRAM WITH ADJUSTED CUT LEVEL TO MOVE ONE HIT INTO THE RED CLUSTER.	329
FIGURE 8.23. PRIMARY HYDROGEN BONDING MOTIF EXHIBITED BY ALL MEMBERS OF THE RED CLUSTER ILLUSTRATED BY BUDWAY10 [160].	329
FIGURE 8.24. SINGLE HYDROGEN BOND GIVING RISE TO THE YELLOW CLUSTER MILFAP [162].	330
FIGURE 8.25. SIMILAR HYDROGEN BONDING MOTIF IN THE HIT FROM THE GREEN CLUSTER; BUDWAY10 [160], LEFT AND THE TWO MEMBERS OF THE CYAN CLUSTER, MIDDLE EGMCYT10 [161] AND RIGHT MILFAP [162].	331
FIGURE 8.26. DENDROGRAM AFTER SECOND MOVEMENT OF THE SIMILARITY CUT-OFF TO INCORPORATE THE CYAN AND GREEN CLUSTERS INTO ONE GREEN CLUSTER.	331
FIGURE 8.27. SEARCH FRAGMENT DEFINED FOR CYTOSINE MOLECULAR COMPLEXES IN THE CSD.	332
FIGURE 8.28. HETERODIMER HYDROGEN BONDED RING FORMED ONCE A HYDROGEN ATOM IS TRANSFERRED TO THE CYTOSINE MOLECULE ILLUSTRATED BY CASCIJ [173].	333
FIGURE 8.29. PSEUDO-WATSON-CRICK BASE PAIR WHICH CAN ONLY BE FORMED IN THE PRESENCE OF BOTH NEUTRAL AND PROTONATED CYTOSINE MOLECULES SHOWN BY CYTZNC [174].	334
FIGURE 8.30. PSEUDO-WATSON-CRICK BASE PAIRS LINKED VIA TYPE B BONDING MOTIFS REPRESENTED BY CYTRES10 [169].	334
FIGURE 8.31. TWO HYDROGEN ATOMS PRESENT EITHER SIDE OF AN INVERSION CENTRE WITH EACH HAVING AN OCCUPANCY FACTOR OF 0.5 REPRESENTED BY ODICOU [163].	335
FIGURE 8.32. TYPE C HYDROGEN BONDING MOTIF FOUND IN SEVERAL OF THE CSD EXAMPLES ILLUSTRATED BY CYTOSM [170].	336
FIGURE 8.33. MOST COMMON STATE FOR A CYTOSINE MOLECULE TO EXIST IN THE CSD REPRESENTED BY ADALAS [171].	337

Chapter 9

FIGURE 9.1. HYDROGEN BONDED RING MOTIF EXHIBITED BY THE MOLECULAR COMPLEX OF CYTOSINIUM 5-FLUOROSALICYLATE (LEFT) AND THE SUPRAMOLECULAR UNIT EXHIBITED BY THE CYTOSINE AND 2,4-DIHYDROXYBENZOIC ACID HYDRATE MOLECULAR COMPLEX (RIGHT).....	343
FIGURE 9.2. COMMON BONDING PATTERN WITH ONLY SINGLE HYDROGEN BONDS LINKING HYDROGEN BONDED RING UNITS TOGETHER EXHIBITED BY THE MOLECULAR COMPLEX OF CYTOSINIUM 5-FLUOROSALICYLATE.	343
FIGURE 9.3. PSEUDO-WATSON-CRICK BASE PAIR MOTIF EXHIBITED BY THE MOLECULAR COMPLEX OF CYTOSINE AND 4-FLUOROBENZOIC ACID HYDRATE.....	344
FIGURE 9.4. TWO POSSIBLE BONDING PATTERNS – EXTENDED CYTOSINE HOMO-BASE-PAIRED CHAIN AS EXHIBITED BY THE MOLECULAR COMPLEX OF CYTOSINE AND 2-FLUOROBENZOIC ACID HYDRATE (LEFT) AND THE FORMATION OF A SUPRAMOLECULAR UNIT AS SEEN IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 3,5-DINITROBENZOIC ACID (RIGHT).	344
FIGURE 9.5. THE EXTENDED HOMO BASE-PAIRED CHAIN FORMED BY ALTERNATING TYPE A AND TYPE B BONDING MOTIFS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 3-FLUOROBENZOIC ACID SOLVATE (LEFT) AND THE TERMINATED CHAIN SEEN IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 4-HYDROXYBENZOIC ACID.....	345
FIGURE 9.6. SEPARATE CHAINS LINKED VIA WEAK HYDROGEN BONDS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND 4-FLUOROBENZOIC ACID (GREEN REPRESENTS THE HYDROGEN BONDS THAT MAKE USE OF THE FLUORINE ATOMS AND THE RED AND BLACK REPRESENT SEPARATE CHAINS).....	346
FIGURE 9.7. SPACEFILL REPRESENTATION TO SHOW THE VOIDS PRESENT IN THE MOLECULAR COMPLEXES OF 3-FLUOROBENZOIC ACID AND 5-FLUOROCYTOSINE (LEFT) AND 4-FLUOROBENZOIC ACID AND 5-FLUOROCYTOSINE (RIGHT).	346
FIGURE 9.8. SINGLE F···H-C HYDROGEN BONDS CONNECTING CHAINS IN THE STRUCTURES OF 5-FLUOROCYTOSINE AND BENZOIC ACID (LEFT) AND 5-FLUOROCYTOSINE AND 2-FLUOROBENZOIC ACID (RIGHT).....	347
FIGURE 9.9. FLUORINE INTERACTIONS IN LEFT, CYTOSINIUM 5-FLUOROSALICYLATE AND RIGHT 5-FLUOROCYTOSINIUM SALICYLATE MOLECULAR COMPLEXES.	348
FIGURE 9.10. BASE-PAIR STYLE MOTIF INVOLVING FLUORINE HYDROGEN BONDS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINIUM 2,6-DIHYDROXYBENZOATE HYDRATE FORM I.....	349
FIGURE 9.11. FLUORINE INTERACTIONS LINKING SEPARATE CHAINS IN THE MOLECULAR COMPLEX OF 5-FLUOROCYTOSINE AND UREA.....	350
FIGURE 9.12. OVERLAID MOLECULAR COMPLEXES OF URACIL AND UREA (RED) AND 5-FLUOROURACIL AND UREA (GREEN) SHOWING SUBTLE INTERACTIONS DUE TO THE FLUORINE ATOM. LEFT SHOWS THE TWO STRUCTURES MOVING OUT OF SYNC AND RIGHT SHOWS THE FLUORINE INTERACTIONS WHICH CAUSE THIS SLIGHT CHANGE IN ORIENTATION.	351

FIGURE 9.13. STACKING INTERACTIONS IN THE URACIL AND UREA MOLECULAR COMPLEX (RED REPRESENTS URACIL AND GREEN REPRESENTS UREA).....	351
FIGURE 9.14. HOMO BASE-PAIRING (LEFT) AND. ALTERNATING PAIRS OF 5- FLUOROURACIL AND THIOUREA MOLECULES IN THE HYDROGEN BONDED CHAIN	352
FIGURE 9.15. TOP, BASE PAIRING MOTIF IN THE MOLECULAR COMPLEX OF 5- FLUOROURACIL 3,5-DIHYDROXYBENZOIC ACID HYDRATE INVOLVING C-H...F HYDROGEN BONDS. BOTTOM, WATER MOLECULES SPACING THE CONVENTIONAL BASE PAIR.....	353

Table of tables

TABLE 1.1. THE DEFINITION AND CHARACTERISTICS OF HYDROGEN BONDS [TAKEN FROM REFERENCE 31].	47
TABLE 2.1 DIFFRACTOMETERS USED AND ASSOCIATED SOFTWARE USED ...	94
TABLE 3.1. THE HYDROGEN BONDS FOUND WITHIN A CONVENTIONAL WATSON-CRICK UNIT OF THE 9-ETHYLGUANINE 1-METHYLCYTOSINE MOLECULAR COMPLEX (EGMCYT10 [131]).	107
TABLE 3.2. THE HYDROGEN BONDS FOUND WITHIN THE PSEUDO-WATSON- CRICK CYTOSINE-CYTOSINIUM UNIT OF THE MOLECULAR COMPLEX OF CYTOSINE WITH BENZOIC ACID.	107
TABLE 3.3 BOND DISTANCES IN EACH MOLECULE PRESENT IN THE MOLECULAR COMPLEX OF CYTOSINE AND BENZOIC ACID	109
TABLE 3.4. HYDROGEN BOND DISTANCES OF THE PSEUDO-WATSON-CRICK HYDROGEN BONDING MOTIF IN THE MOLECULAR COMPLEX OF CYTOSINE AND 2-FLUOROBENZOIC ACID HEMIHYDRATE.....	114
TABLE 3.5. HYDROGEN BOND DISTANCES OF THE HYDROGEN BONDED RING IN THE MOLECULAR COMPLEX OF CYTOSINIUM 3-FLUOROBENZOATE.	120
TABLE 3.6. HYDROGEN BOND DISTANCES OF THE PSEUDO-WATSON-CRICK HYDROGEN BONDING MOTIF IN THE MOLECULAR COMPLEX OF CYTOSINE AND 4-FLUOROBENZOIC ACID HYDRATE.....	123
TABLE 3.7 THE ΔPK_A VALUES, LEVEL OF HYDROGEN TRANSFER AND THE RATIO OF THE COMPONENTS IN THE FOUR MOLECULAR COMPLEXES	127
TABLE 3.8 CRYSTALLISATION RATIO, NUMBER OF MOLECULAR SPECIES AND PRESENCE OF WATER FOR THE FOUR STRUCTURES	128
TABLE 3.9 COMPARISON OF THE THREE PSEUDO-WATSON-CRICK BONDING MOTIFS SEEN IN THE MOLECULAR STRUCTURES OF CYTOSINE WITH BENZOIC ACID AND 2- AND 4-FLUOROBENZOIC ACID.....	130
TABLE 3.10. HYDROGEN BOND DISTANCES CONNECTING THE PSEUDO- WATSON-CRICK BASE PAIRS INTO CYTOSINE CHAINS.	130
TABLE 3.11. SUMMARY OF CRYSTALLOGRAPHIC DATA FOR THE MOLECULAR COMPLEXES OF CYTOSINE WITH BENZOIC ACID AND 2-,3- AND 4-FLUOROBENZOIC ACID.....	136
TABLE 3.12 HYDROGEN BONDS IN THE MOLECULAR COMPLEX OF 5- FLUOROCYTOSINE AND BENZOIC ACID WITH HYDROGEN BOND DISTANCES COLOUR CODED WITH FIGURE 3.52.....	139
TABLE 3.13. HYDROGEN BONDS IN THE MOLECULAR COMPLEX OF 5- FLUOROCYTOSINE AND 2-FLUOROBENZOIC ACID WITH BOND DISTANCES COLOUR CODED WITH FIGURE 3.61.....	143
TABLE 3.14 CRYSTAL STRUCTURE INFORMATION FOR THE MOLECULAR COMPLEXES OF 5-FLUOROCYTOSINE WITH BENZOIC ACID AND 2-,3- AND 4-FLUOROBENZOIC ACID.....	157
TABLE 3.15. ΔPK_A VALUES, LEVEL OF HYDROGEN TRANSFER AND THE CRYSTALLISATION RATIO FOR THE MOLECULAR COMPLEXES OF 5- FLUOROCYTOSINE WITH BENZOIC ACID AND 2-, 3- AND 4- FLUOROBENZOIC ACID.....	158
TABLE 3.16. ANGLE OF THE BENZOIC ACID MOLECULE WITH RESPECT TO THE 5-FLUOROCYTOSINE HOMO-BASE PAIRED CHAINS FOR THE MOLECULAR COMPLEXES OF 5-FLUOROCYTOSINE WITH BENZOIC ACID AND 2-,3- AND 4-FLUOROBENZOIC ACID	159

TABLE 3.17. COMPARISON OF THE BASE STACKING INTERACTIONS FOR THE MOLECULAR COMPLEXES OF 5-FLUOROCYTOSINE WITH BENZOIC ACID AND 2-,3- AND 4-FLUOROBENZOIC ACID.	161
TABLE 3.18. % OF THE UNIT CELL MADE UP OF THE VOID IN THE MOLECULAR COMPLEXES OF 5-FLUOROCYTOSINE WITH 3-FLUOROBENZOIC ACID AND 4-FLUOROBENZOIC ACID.....	164
TABLE 4.1. SPACING BETWEEN LAYERS IN THE CYTOSINIUM SALICYLATE MOLECULAR COMPLEX.....	171
TABLE 4.2. CRYSTALLOGRAPHIC DATA FOR THE MOLECULAR COMPLEXES OF CYTOSINE AND 5-FLUOROCYTOSINE WITH SALICYLIC ACID AND 5-FLUOROSALICYLIC ACID	172
TABLE 4.3 COMPARISON OF THE SPACING BETWEEN LAYERS IN THE CYTOSINIUM SALICYLATE, 5-FLUOROSALICYLATE AND CYTOSINIUM 5-FLUOROSALICYLATE MOLECULAR COMPLEXES.	180
TABLE 4.4 ANGLE CREATED AT CORNER OF L SHAPE CREATED IN THE THREE RELATED STRUCTURES.	181
TABLE 4.5. HYDROGEN BOND DISTANCES OF THE PSEUDO-WATSON-CRICK HYDROGEN BONDING MOTIF IN THE 5-FLUOROCYTOSINE AND 5-FLUOROSALICYLIC ACID HYDRATE MOLECULAR COMPLEX. THE NUMBERING REFERS TO THAT SHOWN IN FIGURE 4.21.....	183
TABLE 4.6 HYDROGEN BOND DISTANCES OF THE PSEUDO-WATSON-CRICK HYDROGEN BONDING MOTIF IN THE CYTOSINE 5-FLUOROSALICYLIC ACID HYDRATE MOLECULAR COMPLEX.....	188
TABLE 4.7. DISTANCES BETWEEN PLANES IN THE MOLECULAR COMPLEX OF CYTOSINE 5-FLUOROSALICYLIC ACID HYDRATE.....	191
TABLE 4.8. HYDROGEN BONDS OF THE PRIMARY BONDING MOTIF IN DIMER 1 AND 2.	194
TABLE 4.9. CRYSTALLOGRAPHIC DATA FOR THE MOLECULAR COMPLEXES OF CYTOSINE AND 5-FLUOROCYTOSINE WITH 3 AND 4-HYDROXYBENZOIC ACID.....	198
TABLE 4.10. HYDROGEN BOND DISTANCES FOR THE PSEUDO-WATSON-CRICK BONDING MOTIF IN THE CYTOSINE 4-HYDROXYBENZOIC ACID HYDRATE MOLECULAR COMPLEX.....	203
TABLE 4.11. HYDROGEN BONDS INVOLVING THE PSEUDO-WATSON-CRICK BASE PAIR IN 4-HYDROXYBENZOIC ACID AND CYTOSINE HYDRATE MOLECULAR STRUCTURE.	204
TABLE 4.12 THE ΔpK_A AND THE % TRANSFER OF A PROTON WITH RESPECT TO THE CYTOSINE MOLECULES	210
TABLE 4.13. HYDROGEN BONDS OF THE PRIMARY HYDROGEN BONDING MOTIF, THE HETERODIMER, IN THE FULLY CHARGED MOLECULAR COMPLEXES.	214
TABLE 4.14 HYDROGEN BONDS INVOLVING THE PRIMARY HYDROGEN BONDING MOTIF IN THE THREE ANHYDROUS FULLY CHARGED MOLECULAR COMPLEXES.	215
TABLE 4.15. UNIT CELLS OF THE THREE ANHYDROUS FULLY CHARGED MOLECULAR COMPLEXES.	216
TABLE 4.16. HYDROGEN BONDS OF THE PSEUDO-WATSON-CRICK HYDROGEN BONDING MOTIF	218
TABLE 4.17. HYDROGEN BONDING IN THE ISOSTRUCTURES OF CYTOSINE AND 5-FLUOROSALICYLIC ACID HYDRATE AND 5-FLUOROCYTOSINE AND 5-FLUOROSALICYLIC ACID HYDRATE.....	220

TABLE 5.1. THE HYDROGEN BONDS INVOLVING THE 5-FLUOROURACIL MOLECULE IN THE 1:1 MOLECULAR COMPLEX WITH UREA. A-E REPRESENT THE LABELLING IN FIGURE 5.1.	224
TABLE 5.2. THE HYDROGEN BONDS INVOLVING THE 5-FLUOROURACIL MOLECULE IN THE 1:1 MOLECULAR COMPLEX WITH UREA. A-E REPRESENT THE LABELLING IN FIGURE 5.10.	228
TABLE 5.3. THE HYDROGEN BONDS INVOLVING THE 5-FLUOROURACIL AND THIOUREA MOLECULES IN THE 1:1 MOLECULAR COMPLEX OF 5FLUOROURACIL AND THIOUREA. A-E REPRESENT THE LABELLING IN FIGURE 5.19.	233
TABLE 5.4. THE HYDROGEN BONDS INVOLVING THE 5-FLUOROCYTOSINE MOLECULE IN THE 2:1 MOLECULAR COMPLEX WITH UREA. A-D REPRESENT THE LABELLING IN FIGURE 5.36.	243
TABLE 5.5. A SUMMARY OF THE RATIOS OBTAINED IN THE MOLECULAR COMPLEXES INVOLVING UREA AND THIOUREA.....	245
TABLE 5.6. COMPARISON OF THE DISTANCES DEFINE THE BASE PAIR MOTIF IN THE URACIL WITH UREA AND THE 5-FLUOROURACIL AND UREA MOLECULAR COMPLEXES.	246
TABLE 5.7. COMPARISON OF THE DISTANCES IN THE BIFURCATED HYDROGEN BONDS IN THE URACIL WITH UREA AND THE 5-FLUOROURACIL AND UREA MOLECULAR COMPLEXES.....	247
TABLE 5.8. CRYSTALLOGRAPHIC DATA FOR THE MOLECULAR COMPLEXES REPORTED IN CHAPTER 5.....	253
TABLE 6.1. THE HYDROGEN DISTANCES FOR THE HYDROGEN BONDS BETWEEN 5-FLUOROURACIL MOLECULES AND WATER MOLECULES IN THE 5-FLUOROURACIL 3,5-DIHYDROXYBENZOIC ACID HYDRATE MOLECULAR COMPLEX. THE LETTERING CORRESPONDS TO THE HYDROGEN BONDS SHOWN IN FIGURE 6.41.	279
TABLE 6.2. HYDROGEN BOND LENGTHS OF THE 3,5-DIHYDROXYBENZOIC ACID DIMERS AND THE HYDROGEN BONDED RINGS IN THE MOLECULAR COMPLEX OF 5-FLUOROURACIL 3,5-DIHYDROXYBENZOIC ACID HYDRATE. BOND LABELING CORRESPONDS TO THOSE INDICATED IN FIGURE 6.44.	281
TABLE 6.3. THE HYDROGEN BOND DISTANCES IN THE PSEUDO-WATSON-CRICK UNIT IN THE 5-FLUOROCYTOSINE 3,5-DINITROBENZOIC ACID MOLECULAR COMPLEX.....	285
TABLE 6.4. SUMMARY OF THE MOLECULAR COMPLEXES OF CYTOSINE AND 5-FLUOROCYTOSINE REPORTED IN THIS CHAPTER.....	287
TABLE 6.5. CRYSTALLOGRAPHIC DATA FOR THE MOLECULAR COMPLEXES REPORTED IN CHAPTER 6.....	291
TABLE 7.1 HYDROGEN BONDS OF THE PSEUDO-WATSON-CRICK BASE PAIR IN THE MOLECULAR COMPLEX OF CYTOSINE ISONICOTINIC ACID HYDRATE.....	293
TABLE 7.2 HYDROGEN BONDS FOR THE PSEUDO-WATSON-CRICK HYDROGEN BOND MOTIF IN THE MOLECULAR COMPLEX OF CYTOSINE AND 2-NITROBENZOIC ACID	300
TABLE 7.3. THE THREE HYDROGEN BONDS IN THE PSEUDO-WATSON-CRICK BASE-PAIRING MOTIF IN THE MOLECULAR COMPLEX OF CYTOSINE AND 3-BROMOBENZOIC ACID.....	305
TABLE 7.4 RATIO OF PRODUCTS, PERCENTAGE OF CYTOSINE MOLECULES GAINING A HYDROGEN AND Δ PKA VALUES FOR THE STRUCTURES IN CHAPTER 7	310
TABLE 7.5 PKA VALUES OF THE STARTING MATERIALS.....	310

TABLE 7.6. TABLE OF CRYSTALLOGRAPHIC DATA FOR THE MOLECULAR COMPLEXES OF CYTOSINE IN CHAPTER 7.....	311
TABLE 8.1 – HYDROGEN BONDING MOTIFS IN THE DSNAP ANALYSIS OF ADENINE THYMINE BASE PAIRS.	317
TABLE 8.2. HYDROGEN BOND DISTANCES BETWEEN THE BASE PAIRS IN THE RED, YELLOW AND GREEN CLUSTERS.....	319
TABLE 8.3. HYDROGEN BOND DISTANCES BETWEEN THE BASE PAIRS IN THE CYAN CLUSTER.....	321
TABLE 8.4. HYDROGEN BOND DISTANCES FOR THE REVERSE WATSON-CRICK PRIMARY HYDROGEN BONDING MOTIF OF THE BLUE AND GREEN STRIPED CLUSTERS.	322
TABLE 8.5. HYDROGEN BOND DISTANCE OF SINGLE HYDROGEN BOND BETWEEN THE ADENINE AND THYMINE IN PINK CLUSTER.....	323
TABLE 8.6. HYDROGEN BOND DISTANCES IN THE WATSON CRICK BASE PAIR BETWEEN THE ADENINE AND THYMINE IN ORANGE STRIPED CLUSTER.	324
TABLE 8.7. BOND DISTANCES OF HYDROGEN BONDS FOR THE BASE PAIR BETWEEN THE CYTOSINE AND GUANINE IN THE RED CLUSTER. GOING DOWN COLUMN ONE REPRESENTS GOING FROM LEFT TO RIGHT IN THE CLUSTER.....	330
TABLE 8.8. THE HYDROGEN BONDING MOTIFS FOUND IN CHARGED CYTOSINE MOLECULAR COMPLEXES.	333
TABLE 8.9. THE NUMBER OF STRUCTURES IN THE CSD AND THIS THESIS THAT SHOW MIXED CYTOSINE AND CYTOSINIUM MOLECULAR COMPLEXES AND THE HYDROGEN BONDING MOTIFS FOUND.	335
TABLE 8.10. NUMBER OF STRUCTURES FROM CSD AND THESIS THAT CONTAIN NO HYDROGEN TRANSFER AND THE HYDROGEN BONDING MOTIFS FOUND	336
TABLE 8.11. COMPARISON OF THE PREVALENCE OF HYDROGEN TRANSFER IN THE CSD AND THESIS STRUCTURES.	337
TABLE 8.12. COMPARISON OF THE PREVALENCE OF THE HYDROGEN BONDING MOTIF IN THE CSD AND THESIS STRUCTURES.	338
TABLE 9.1 ΔpK_A VALUES FOR THE CYTOSINE OR 5-FLUOROCYTOSINE MOLECULAR COMPLEXES THAT RESULTED IN A 1:1 MOLECULAR COMPLEX RATIO	340
TABLE 9.2 ΔpK_A VALUES FOR THE CYTOSINE OR 5-FLUOROCYTOSINE MOLECULAR COMPLEXES NOT IN 1:1 RATIO AND CONTAINING NO SOLVENT	341
TABLE 9.3 ΔpK_A VALUES FOR THE CYTOSINE OR 5-FLUOROCYTOSINE MOLECULAR COMPLEXES NOT IN 1:1 RATIO BUT CONTAINING SOLVENT	341
TABLE 9.4 TOTAL NUMBER OF CYTOSINE OR 5-FLUOROCYTOSINE MOLECULAR COMPLEXES WITH PROTONATED, MIXED PROTONATED STATES AND NEUTRAL ONLY CYTOSINE MOLECULES.....	342

Chapter 1 – Introduction

The aim of the work presented in this thesis was to investigate the effect of the addition of a fluorine atom to one or both of the components in molecular complexes involving nucleobases on the consequent hydrogen bonding motifs obtained. This builds on recent work on the incorporation of such modified nucleosides into polymeric DNA or DNA-related structures [1,2,3,4]. A range of these substituted nucleosides can also be polymerised using DNA polymerases, allowing these materials to go beyond synthetic chemical methods in producing chemically modified DNA materials. These substituted materials have potential impact in many fields.

1.1 DNA

Deoxyribonucleic acid, or DNA, is the genetic material present in a cell. This genetic material is used in the development and functioning of all living organisms. The main purpose of DNA is to store information for long-term storage. DNA is a linear polymer built up of monomeric units, the nucleotides. The nucleotides are connected with a backbone made of sugars and phosphate groups joined by an ester bond. This forms a chain of nucleotides.

The DNA helical structure is a molecular assembly with a remarkable biological functionality and selectivity in a biological context. The basic molecular assembly of DNA consists of a double helical structure, comprising two intertwined sugar-phosphate backbones, with the two strands linked by complementary pairs of bases (Figure 1.1). These base pairs provide a natural route to the functionality and selectivity of the DNA structure, with hydrogen bond mediated linkages between adenine-thymine and guanine-cytosine pairs. The Wilson group has a long record of studies of hydrogen bonds and indeed of hydrogen-bonded base pairs [5,6,7], and the work presented in this thesis builds on this.

Modifications of the DNA structure can be naturally carried out in two distinct areas: firstly the sugar-phosphate backbone, while retaining its basic function in providing a framework to enable and control base-pairing; secondly, the base-pairing and stacking, and the bases themselves, which are the effective functional groups in the chemical and biological activity of DNA. There has been a recent increase of interest in the possible chemical modification of DNA components, with a view to modifying molecular function while maintaining the critical ability to retain important biological activity. These

investigations must deal with the interplay of factors governing the structure, including the conformation of the sugar-phosphate backbone, the hydrogen-bonded base-pairing (including subtle factors such as propeller-twisting of these bases), the base-stacking (an effect notoriously difficult to quantify fully), and the critical aspect of solvent structure, including strongly hydrogen bonded water networks in either of the two main grooves in the double helical structure.

The information stored in DNA is dependent on the sequence these bases are in. The information is read using the genetic code, which specifies the sequence of the amino acids within proteins. The double helical structure DNA was first proposed by James D. Watson and Francis Crick [8] (Figure 1.1).

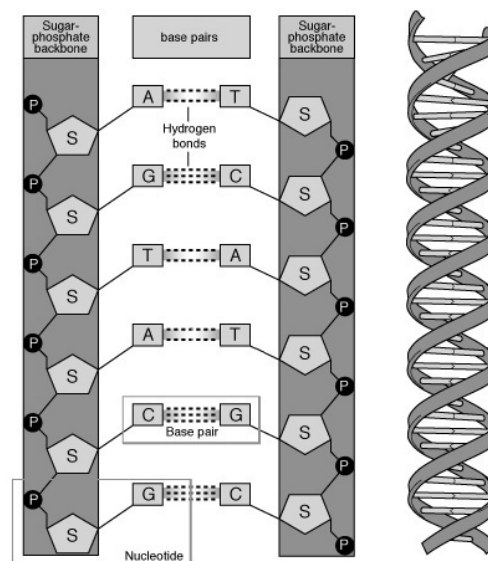


Figure 1.1 Double helical structure of of DNA showing the hydrogen bonds between bases and the sugar phosphate backbone. P represents the phosphate backbone, S the sugar units, and A, T, G, C represent the four nucleobases [9].

Unlike RNA, in all living organisms, DNA does not usually exist as a single strand molecule, but instead as a pair of strands that are held tightly together. The two chains are twisted together to form the shape of a double helix. The nucleotide repeats contain both the segment of the backbone of the molecule, which is what holds the chain together, and a base, which interacts with another DNA strand in the helix. A base linked to a sugar is called a nucleoside and a base linked to a sugar and one or more phosphate groups is called a nucleotide. If multiple nucleotides are linked together, as in DNA, this polymer is called a polynucleotide. The arrangement of two nucleotides binding across the double helix through the bases (via hydrogen bonds) is known as a base pair. The bases in DNA have a very strong selectivity towards pairing in specific complementary pairs (see below); consequence of this base complementary, is that all the information stored in DNA is duplicated on each strand, which is vital in DNA replication.

The code present in DNA is read by copying segments of the related nucleic acid into RNA via a process called transcription. This process creates a complementary RNA copy of a sequence of DNA. RNA and DNA are thus nucleic acids that use base pairs of nucleotides as a complementary language that can be converted back and forth from DNA to RNA by the action of the correct enzyme.

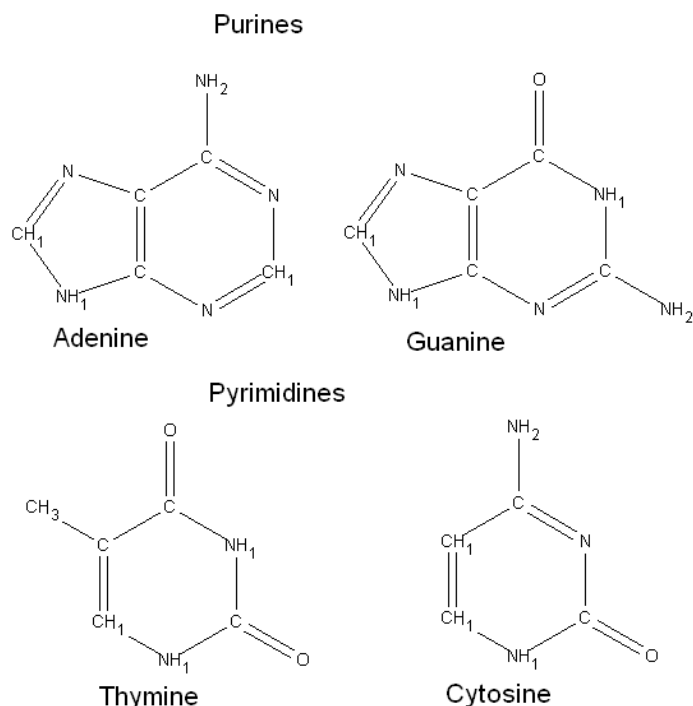


Figure 1.2. The four nucleobases found in DNA, split into two groups, purines (adenine and guanine) and pyrimidines (thymine and cytosine).

There are two main types of bases found in DNA, purines and pyrimidines (Figure 1.2). Each of the two categories of bases contains two bases, the two purines, adenine and guanine, and the two pyrimidines, cytosine and thymine. These bases are highly specific as purines only hydrogen bond or base pair with pyrimidines. Even more specific is the fact that adenine will only base pair with thymine and cytosine will only base pair with guanine as seen in Figure 1.3. In RNA the thymine base is replaced by uracil. The adenine (A) and thymine (T) base pair contains a pair of hydrogen bonds, whilst the guanine (G) and cytosine (C) base pair contains an additional hydrogen bond to make three in the base pair. DNA with a high proportion of GC base pairs is more stable than DNA that has a low content of GC base pairs. This additional stability, is not due to the extra hydrogen bond but rather the contribution of stacking interactions of the base-pairs [10]. The hydrogen bonds provide specificity of the pairing and are not the sole provider of the stability in the double helix.

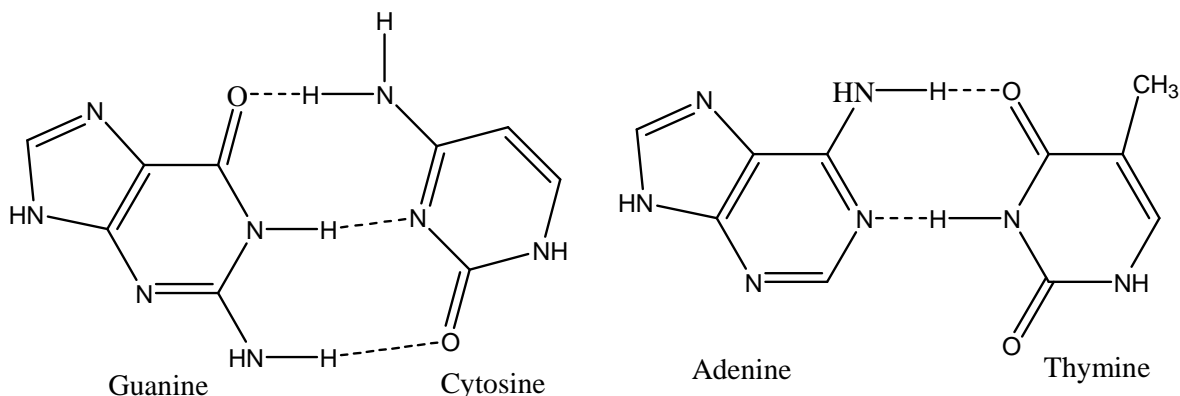


Figure 1.3. The base pairs found in DNA – guanine hydrogen bonding with cytosine and adenine hydrogen bonding with thymine.

It is therefore the percentage of GC pairs and the overall length of the DNA double helix that determines the strength of the association between the strands. As a result long strands with a high GC content have stronger interacting strands when compared with a short strand with a high AT content. Areas of DNA that need to be easily pulled apart, for DNA replication, tend to have a high AT content to make it easier to pull the strands apart. The way to determine the strength of the interactions is to find the temperature required to break the hydrogen bonds present, the melting temperature. When all of the base pairs melt, the two strands separate and exist in solution as two completely independent molecules.

1.2 Nucleosides

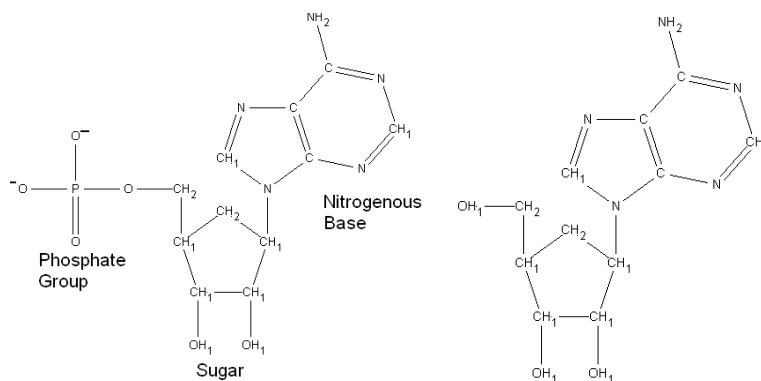


Figure 1.4. The difference between a nucleotide (left – containing a phosphate group) and a nucleoside (right – with no phosphate group).

There are many different functions of nucleosides that can be manipulated for a beneficial biological effect [11,12,13,14]. Nucleosides are glycosylamines that are made by attaching a nucleobase to a ribose or deoxyribose ring (depending on whether it is RNA or DNA). Nucleosides are not to be confused with nucleotides, which have three components: the nitrogenous base, a sugar and a phosphate group. Figure 1.4 shows that the main difference between a nucleoside and a nucleotide is that a nucleoside does not have the phosphate group, instead having a hydroxyl group attached to C5 of the ribose.

Nucleosides have been studied to try and use their activity for antiviral and anticancer treatments. This would be a major breakthrough if these nucleosides could eventually be used for these procedures. Cass *et al* at the University of Alberta, Canada, have investigated the use of nucleosides for antiviral and anticancer treatment [15]. Their research has been into the role of nucleoside transporters in resistance to cancer chemotherapy. The efficacy of these anticancer nucleoside drugs greatly depends on a complex interplay of transporters mediating entry of nucleoside drugs into cells, efflux mechanisms that remove drugs from intracellular compartments and cellular metabolism to active metabolites [11].

1.3 Fluorinated Nucleosides

There has been much investigation carried out into fluorinated nucleosides [16,17,18,19] and fluorinated bases due to the uses that can be attributed to them once fully formulated. This work has been carried out with antiviral and anticancer properties in mind due to the major benefits that could be had if this works and becomes readily available. There has been no major use of these on a global scale as of yet but with the amount of research currently going into this area it may not be a distant reality. The first obstacle to be overcome is the synthesis of the fluorinated nucleotides themselves [20,21]. Many different research groups have carried out work in this area but a group at Harvard (Lee, Uttamapinant and Verdine) have had some success. This group has dealt with the synthesis of 4'-fluoronucleosides [22]. One of the main purposes of these products is to stop cell growth of tumors by inhibiting the replication of the cells. This has not been totally successful so far but more studies have been undertaken and it is hopeful that this process will one day be possible [23,24,25].

Fluorinated nucleosides exhibit a wide variety of biological activity and have been used as anti-tumor and antiviral agents [25]. One factor that is very important is the stability of the nucleoside analogue and in particular the stability of the glycosyl bond in determining the biological activity as well as the therapeutic usefulness of nucleosides as drug candidates. Fluorine can be substituted at the 2'- or 3'- position of a sugar and this is known to increase the chemical stability of the nucleoside analogues, particularly in an acidic environment. This substitution of fluorine also increases the metabolic stability. These findings have suggested the importance of the fluorine moiety in the nucleoside and therapeutic agents [26].

1.4 Crystal Engineering

Crystal engineering [27,28] is the design and in some cases the synthesis of molecular complexes that exhibit a motif that was designed or predicted prior to the molecular assembly. The aims are to adapt the physical and chemical properties of the resulting product. One way of helping to engineer crystal structures is to take into account the bonding motifs, for example the hydrogen bonding, that will be likely to form between two (or more) targeted molecules. The idea is to use functional groups on the same molecule or on separate molecules that will form interactions that are preferred to the other possible interactions. The knowledge of the preferred intermolecular bonding motif when specific functional groups are used can then be used to design alternative molecular complexes. The predictability of crystal structures is the first step towards tuning of the properties. The easiest method for designing and controlling molecular assembly is to target systems where the crystal structure is robustly held together via strong intermolecular interactions such as hydrogen bonding.

The problem with crystal engineering is that a basic knowledge of preferred bonding patterns must be known before predictions and design can be undertaken. This means a process must be adopted to build up a library of basic bonding motifs and a knowledge of preferred bonding patterns when specific functional groups are present. Crystal engineering requires account to be taken of both the strong and weak intermolecular interactions that can occur in the structure under investigation. The structural motifs can be used as building blocks for the extension of the crystal structure; these basic building blocks are referred to as supramolecular synthons. The structural motifs that are adopted by the crystal engineer are ideally repeatedly reoccurring in related molecular complexes and show the same interactions of particular molecules of function groups; an example of this is the interactions that occur between the base molecules found in DNA. The hydrogen bonding seen in DNA is an example of predictability of the type being explored by crystal engineers.

Some important papers regarding crystal engineering have been published in recent times [29]. One example of the applications of crystal engineering is the molecular complexes formed between chloranilic acid (C) and the three picoline isomers (P) in both 1:1 and 1:2 ratios [30]. The important part here is that in several structures it can be seen that common bonding patterns are repeated with supramolecular units forming in the P:C:C:P (1:1 ratio) or a P:C:P (1:2 ratio) unit. The supramolecular units revolve around bifurcated hydrogen bonds involving hydrogen transfer from the chloranilic acid molecule to the picoline and

the predictability of this forming between the targeted molecules (Figure 1.5). The chloranilic acid molecule is symmetrical and as a result means the same hydrogen bonding opportunities are available on both sides of the molecules (Figure 1.6).

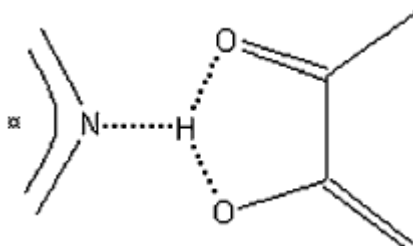


Figure 1.5. Bifurcated hydrogen-bond seen in picoline and chloranilic acid complexes [30].

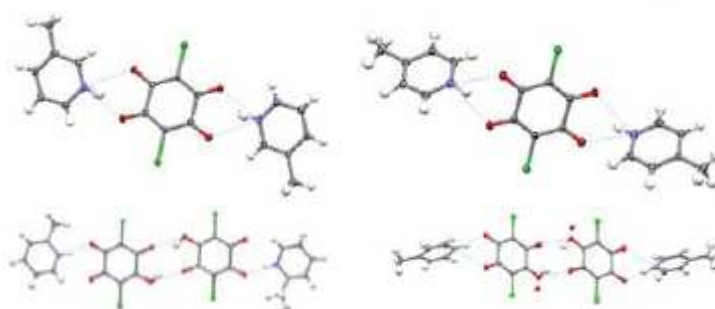


Figure 1.6. P:C:P and P:C:C:P complexes seen in picoline and chloranilic acid complexes [30].

The basic hydrogen bonding of the chloranilic acid and the picoline molecule is readily predicted and this preferred bonding pattern can be seen in both the 1:1 and the 1:2 molecular complexes. The level of proton transfer with respect to the picoline molecule is the same but the chloranilic acid molecule can either transfer both protons in the case of the 1:1 molecular complexes or only one proton in the case of the 1:2 molecular complexes. The protonation state of the chloranilic acid molecule is therefore the structure directing aspect of this particular building block this is another important aspect of the design process in crystal engineering.

1.5 Hydrogen Bonding

A hydrogen bond is, in general, made of three atoms, one hydrogen atom and two electronegative atoms [31]. One electronegative atom will be an acceptor and one will be a donor. These make up a hydrogen bond that can be instrumental to the formation of a molecular complex (Figure 1.7).

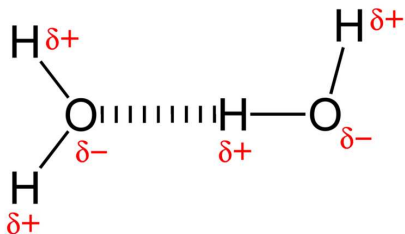


Figure 1.7. Schematic of a conventional hydrogen bond, demonstrated for water.

Table 1.1. The definition and characteristics of hydrogen bonds [taken from reference 31].

	Strong	Moderate	Weak
D-H...A	Mainly Covalent	Mainly Electrostatic	Electrostatic
Bond Energy (kJ mol⁻¹)	60-120	16-60	<12
H...A	1.2-1.5	1.5-2.2	2.2-3.2
D...A	2.2-2.5	2.5-3.2	3.2-4.0
Bond Angles (°)	175-180	130-180	90-150
Relative IR vibration shift (stretching symmetrical mode, cm⁻¹)	25%	10-25%	<10%
¹H NMR chemical shift downfield (ppm)	14-22	<14	?
Examples	Gas phase dimers with strong acids/bases Proton sponge HF complexes	Acids Alcohols Biological molecules	Minor components of bifurcated bonds C-H hydrogen bonds O-H... π hydrogen bonds

For a hydrogen bond interaction to form, the donor must have significant electronegativity to remove electron density from the hydrogen. This will allow the hydrogen to interact with nearby atoms containing lone pairs and polarisable π electrons to form the hydrogen bond as the hydrogen is deshielded. There are three classes of hydrogen bonds: weak, moderate and strong. DNA is one of the archetypal examples of the importance of hydrogen bonds.

1.5.1 Strong Hydrogen Bonds

Strong hydrogen bonds are comparable to covalent bonds and the D...A distance for these bonds is usually less than 2.5 Å [31]. These bonds readily form when there is a deficiency of electron density in the donor group or an excess of electron density in the acceptor group. Due to a deficiency of electrons in the donor group, the resultant positive charge on the hydrogen atom is increased, thus making it more likely to interact. Hydrogen bond formation is greatly increased when there is a surplus of electrons in the acceptor group able to interact with the deshielded proton. It can be possible for neutral donor and acceptor groups to form strong hydrogen bonding interactions. This type of hydrogen bonding also occurs when groups are forced by configuration or conformation into much closer contacts. When these interactions occur they are usually referred to as forced strong hydrogen bonds.

1.5.2 Moderate Hydrogen Bonds

Moderate hydrogen bonds are the most prominent of the three types of hydrogen bond possible [31]. These hydrogen bonds can be inter or intramolecular, and normally consist of neutral donor and acceptor groups. Usually, the donor atoms (D) are more electronegative than the hydrogen and the acceptor atom (A) has a lone pair. An example of this can be the self-association of carboxylic acids or amide interactions in proteins [32,33]. Moderate strength hydrogen bonds do not generally have linear geometry but are often slightly bent.

1.5.3 Weak Hydrogen Bonds

In this case if either or both of the electronegative donor and acceptor atoms are of medium or low electronegativity, then the interaction is classed as a weak hydrogen bond [31]. This situation is normally prominent when the bonded atom is more electroneutral than the H atom. Examples of donor groups are C-H or Si-H. The criteria can also be filled if the acceptor group does not contain a lone pair but instead has π electrons such as an aromatic ring or a carbon-carbon triple bond. A weak hydrogen bond is close to a van der Waals interaction. Weak hydrogen bonds are also less directional than strong and moderate hydrogen bonds. Weak interactions are increasingly being utilised in crystal engineering.

1.5.3.1 Fluorine Hydrogen Bonds

Fluorine hydrogen bonds have been widely recognised since the middle of the last century [34], and in recent years the hydrogen bonding ability of all of the halogens have been investigated. It has been widely seen that fluorine can act as a better hydrogen bond acceptor when compared to the other halogens [34], however, it is not as strong as oxygen and nitrogen [35,36,37,38].

Short contacts between fluorine and hydrogen bound to carbon ($\text{C-F}\cdots\text{H-C}$) represent the main category of interactions involving F atoms. However, short contacts between F and the acidic hydrogen atoms of HO and HN do occur but are rare in the CSD [39]. The CSD is the Cambridge Structural Database, which contains structural information, chemical and bibliographic data for about 400000 organic and metal-organic compounds. Despite the low number of $\text{F}\cdots\text{H}$ contacts when referring to acidic protons there is a statistically significant increase in the short contacts to $\text{C}(\text{sp}^3)\text{-F}$ over $\text{C}(\text{sp}^2)\text{-F}$ bound fluorine atoms.

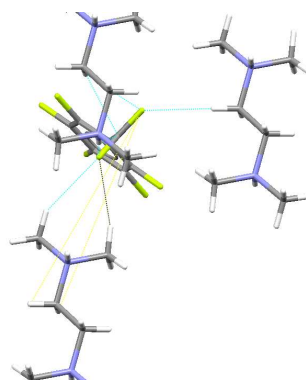


Figure 1.8. Weak strength $\text{F}\cdots\text{H-C}$ hydrogen bond in the molecular complex of N,N,N',N'-tetramethylethylenediammonium bis(difluoro(pentafluorophenyl)methanide) [40].

There are many structures present in the CSD that contain fluorine hydrogen bonds. One such structure was published in 2004 by Zhu [40]. This structure is of N,N,N',N'-tetramethylethylenediammonium bis(difluoro(pentafluorophenyl)-methanide) (Figure 1.8). There are several fluorine interactions in this structure, however, the most relevant interaction is the weak strength hydrogen bond of type $\text{C-H}\cdots\text{F}$. This hydrogen bond has $\text{C}\cdots\text{F}$ distance of 3.384 Å. This interaction can be clearly seen in Figure 1.48 below. This interaction depicts the most common type of hydrogen bond that involves a fluorine atom.

1.5.4 Bifurcated Hydrogen Bonds

A hydrogen bond can be referred to as a bifurcated hydrogen bond when two acceptors or donors are involved. Bifurcated hydrogen bonds have been seen in many structures involving the simplest molecules such as water [41] molecules, and more complex systems such as DNA [42]. The most common type of bifurcated hydrogen bonding involves one donor atom, a single hydrogen atom and two acceptor atoms as shown in Figure 1.9.

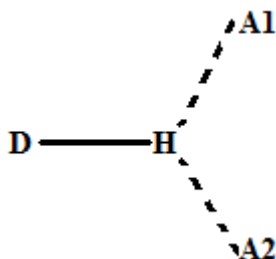


Figure 1.9. The most common type of bifurcated hydrogen bond.

The bifurcated hydrogen bond has some constraints; the hydrogen atom lies in a plane created by the two acceptor atoms and the donor atom. This means the hydrogen atom will lie approximately coplanar to the three atoms making up the plane. Hence, the three angles, one between each bond shown in Figure 1.9, should sum to approximately 360° . The bifurcated hydrogen bond usually has a major component and a minor component. These components mean that one of the acceptor atoms can be more dominant with respect to the other acceptor atom with a shorter $H\cdots A$ distance and a larger $D-H\cdots A$ angle when compared with that of the minor component. It is common for the major component to be directly comparable to a moderate strength hydrogen bond, whilst the minor component is comparable to a moderate or weak strength hydrogen bond. Bifurcated hydrogen bonds are very varied but also very common and can result in varied bond lengths and angles and in some cases proton disorder.

1.5.5 Charge Assisted Hydrogen Bonds

There are two cases when a charge assisted hydrogen bond can exist [43]. The first is a negative charge assisted hydrogen bond. This type of bond is formed then a donor atom has a significant localised negative charge, which results in a strong electrostatic attraction to the hydrogen. The other charge assisted hydrogen bond is a positive charge assisted hydrogen bond. This bond forms when the acceptor hydrogen has a significant localised positive charge, which results in a strong electrostatic attraction to the donor atom. Much work has been done on classifying charge assisted hydrogen bonds [11].

1.5.6 Proton Disorder

In many cases a one-dimensional representation can be constructed to help illustrate the energy minimum in a covalently bonded atom pair with the parent atom at the origin (Figure 1.10).

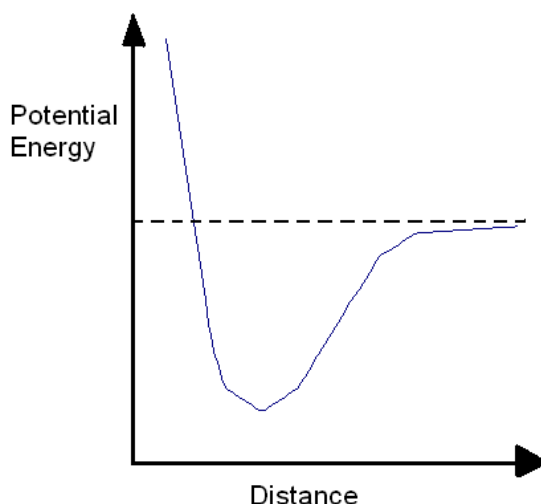


Figure 1.10. Potential energy profile for an atom bonded to another located at the origin of the plot.

Close to the parent atom to which the selected subject atom is bonded there will be a large repulsive force and as a result a large potential. However, in the region of the ideal bond length between the selected atoms is a minimum and as the subject atom is moved further away the potential converges on the value for the two atoms being dissociated. These potential energy wells can become of great use to represent the probable density of hydrogen in a hydrogen bonds [44]. The minimum present in the potential energy well corresponds to the equilibrium position of the hydrogen. The main difference between this and a hydrogen atom in a hydrogen bond is that the hydrogen bond will interact with two atoms, hence there are two such wells in opposite orientations in this representation. Figure 1.11 helps to illustrate three possible hydrogen bond potential energy wells, an unsymmetrical double minimum, a symmetrical double well, a flat-bottomed single well.

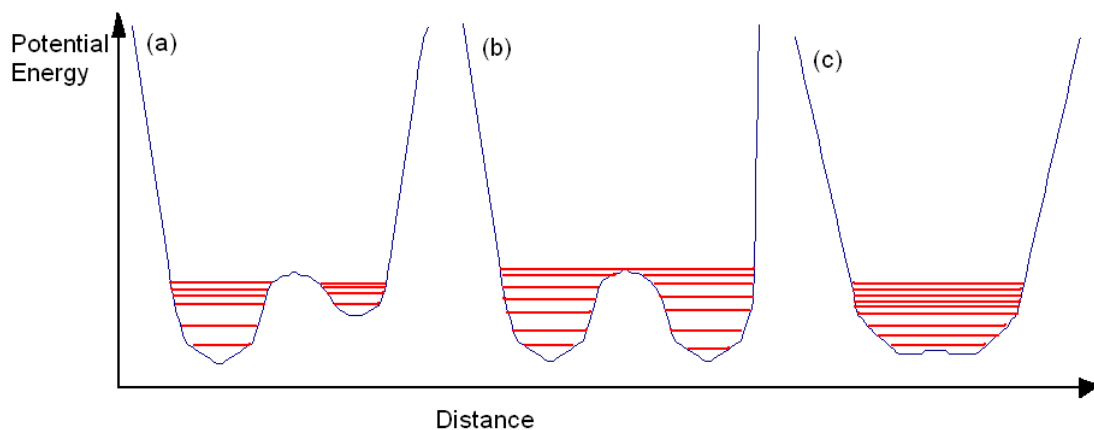


Figure 1.11. Three different types of hydrogen bond potential: a) unsymmetrical double minimum, b) symmetrical double well, c) flat bottom single well.

The first example (unsymmetrical double minimum) is the normal type for a moderate hydrogen bond. There is a second high energy minimum closer to the acceptor atom but this will not normally be occupied by the hydrogen atom. The second example is simply a special case of the first example, with two minima that are equal. The hydrogen can be located in both minima with equal probability i.e. disordered. However, in some environments these two positions can be close in energy but not identical and both can be partly occupied. The final example arises when the two wells of the double well potential merge and the barrier height becomes very low, which is often the case in short, strong hydrogen bonds. This means that the hydrogen can transfer between the donor and the acceptor with only a small increase in energy. This type often shows so-called proton migration as the equilibrium position for the hydrogen can move towards the middle of the potential well with energy gained from the increased temperature [45,46].

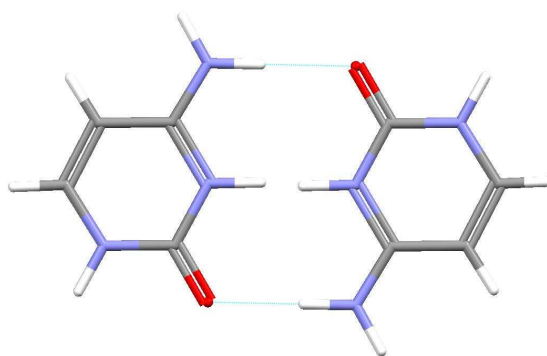


Figure 1.12. An example of a disordered hydrogen atom within a hydrogen bond [47].

The molecular complex of cytosine with 2,5-diethyl-7,7,8,8-tetracyanoquino-dimethane [47] contains a hydrogen bonded unit which shows a disordered hydrogen atom within a hydrogen bond (Figure 1.12). This structure has two cytosine molecules that form what appears like a Watson-Crick hydrogen bonding motif with three hydrogen bonds. However, the hydrogen bond in the middle has two positions for the hydrogen bond with a

50:50 occupancy of each position. This disorder is at its simplest level as the hydrogen atom is disordered by symmetry. In some cases, the proportion of occupancy of each of the two sites is different, for example the case of disorder in carboxylic acid dimers such as that reported for benzoic acid [48]. In this case, the proportion of disorder is found to change as a function of temperature.

1.6 Hydrogen Transfer and pK_a matching

The prediction of whether or not there will be proton transfer between molecular components in a molecular complex is another challenge in crystal engineering. One method of rating the possibility of hydrogen transfer is to take into consideration the pK_a values of the hydrogen bonded molecules used in a co-crystallisation attempt. When the pK_a values of two molecules have a big difference then the proton would be expected to be bound to the molecule with the higher pK_a value. When the pK_a values of the two targeted molecules are within close proximity or are of equal value the problem becomes more complex.

The rule that will be adopted during this thesis was published by Childs *et al* in 2007 [49]. The rule stated here is that when the difference between the pK_a s of the two components is less than 0 it would be expected that the product produced would be a co-crystal with neutral products. If the difference in pK_a was greater than 3.0 then a salt would be expected to form. However, as stated previously in the literature [49] the region of 0-3 is difficult to determine and that this method is inappropriate for the formation of a salt as either salt or a co-crystal can form in this region. If the difference between the two pK_a values is between 0 and 3.00 then a continuum is expected where proton transfer may or may not occur. These rules have been shown to be a helpful guide in the design of molecular complexes [50], however there are examples of complexes that do not follow these rules [51].

Another factor in the success of pK_a matching is the crystallization ratio of the crystalline product. Some work has been carried out that help to investigate the applicability of pK_a matching to molecular complexes that are no longer formed in a 1:1 ratio [52]. This work investigated the co-crystallisation of pentachlorophenol (PCP) with a series of dimethylpyridines (lutidines) and resulted in the formation of molecular complexes that showed strong hydrogen bonding in the solid state. However, competition for the hydrogen atom in these strong hydrogen bonds by the pK_a matched molecules led to a variable degree of hydrogen transfer from PCP towards lutidine which depended on the difference of the pK_a values in the particular complex. There was a strong correlation found between

the ΔpK_a values and the hydrogen transfer when referring to the 1:1 molecular complexes. However, when the crystallisation ratio changed or when a solvent molecule was introduced, the pK_a matching did not work well. The authors state that this occurrence has been widely reported before and is not unexpected but it is relevant in areas such as the pharmaceutical industry towards understanding the crystallisation behaviour in molecular complexes.

1.7 Halogen Bonding and Fluorine Interactions.

In recent years much research has been applied to the role of halogens and their interactions within crystal structures. Crystal engineering has begun to utilise interactions involving halogen atoms, the interactions of which are not as strong or as controllable as hydrogen bonding, but which are still directional and structure directing in nature.

Significant intermolecular interactions in crystal structures can be defined by those that have distances that are shorter than the sum of the van der Waals radii of the two atoms. With all halogen interactions there is a directional preference with which contacting groups position themselves relative to each other. Halogen-halogen contacts can be separated into two groups, type I and type II. This separates the interactions into two preferred geometries, type I ($\theta_1=\theta_2$) and type II ($\theta_1=180^\circ$, $\theta_2=90^\circ$) where θ_1 and θ_2 are the two C-halogen...halogen angles [53].

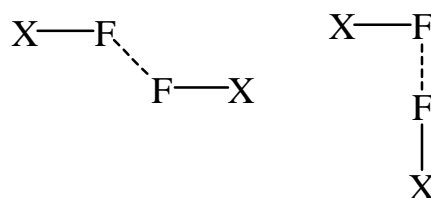


Figure 1.13, Halogen-halogen interactions – type 1 (left) and type 2 (right)

These preferred geometries may result from either:

1. specific attractive forces in certain directions [54,55] such as increased attraction;
2. non-spherical shapes with polar flattening [56,57] in close packed crystals such as decreased repulsion.

In the first of these, the short and directional interactions are caused by attractive forces and this differs from the second one in that it is due to the close packing of non-spherical atomic moieties of molecules.

Desiraju [53] gave an insight into the fact that contacts from F to F, F to H, and F to C are very distinctive. The F...H interactions are present in greater numbers and this is due to the

strong dipole nature of the F \cdots H interaction. This led Desiraju to conclude that the F \cdots F interactions do not have any additional stabilising role in close packing unlike the other halogen-halogen interactions.

The fluorine atom present in many crystal structures can also take part in F \cdots F interactions as well as fluorine hydrogen bonds. In the CSD there is a large proportion of structures that contain the F \cdots F interaction, with the CSD containing more than 700 F \cdots F interactions that are less than double the van der Waals radius of fluorine (2.94 Å). The majority of the F \cdots F interactions fall in the range 2.6 Å to 2.8 Å.

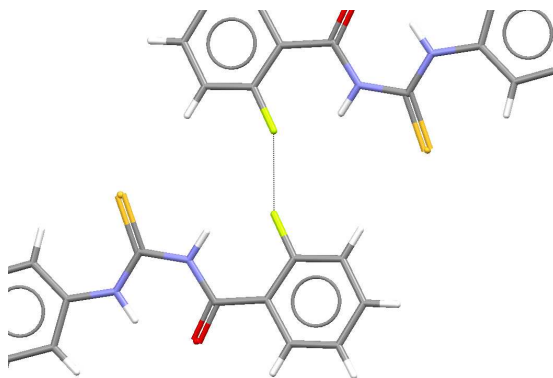


Figure 1.14, F \cdots F interaction present in the structure of N-(2-Fluorobenzoyl)-N'-(p-tolyl)thiourea [58].

In 2004, a structure was published that exhibited a strong F \cdots F interaction, by Zhou *et al* [58], of N-(2-Fluorobenzoyl)-N'-(p-tolyl)thiourea. In this structure there are only weak interactions between separate molecules and as such the fluorine-fluorine interaction plays an important role in linking separate molecules together (Figure 1.14). The fluorine-fluorine interaction is of length 2.722 Å, which is in the middle of region where the majority of fluorine-fluorine interactions fall.

1.8 Other Intermolecular Interactions

1.8.1 π - interactions

π -interactions or π -effects are a non-covalent interaction that involves delocalised π systems. This behaves much like an electrostatic interaction where a region of negative charge interacts with a positive charge. The electron-rich π system can interact with a metal (cationic or neutral), an anion, another molecule and even another π system. These non-covalent interactions are essential to biological events such as protein-ligand recognition.

There are a large number of π interactions that can be present in a wide range of crystal systems. However, there are some that are far more common than others. One of the most

common is metal- π interactions that will involve the interaction of a metal and the face of a nearby π -system. The metal in this case can be a cation or can be neutral. Another interaction that is widely seen is polar- π interactions. This interaction will involve a polar molecule and the quadrupole moment of a π system.

One of the next most common interactions and most important to this research is aromatic-aromatic interactions which can also be known as π stacking or more importantly base-stacking. Base stacking interactions are present in DNA and RNA and are due to dispersion attraction, short-exchange repulsion, and electrostatic interactions, which also contribute to stability. GC over GC stacking interactions are more favourable to the AT over AT stacking interactions, however, it must be taken into account that a GC-GC stacking interaction with the next base pair is geometrically different from a GC-CG interaction. Aromatic-aromatic interactions involve interactions of aromatic molecules with each other. When dealing with base molecules in regards to DNA there are three main stacking interactions that can be undertaken by the molecules: face to face, edge to edge and offset π stacking. It is common in the literature that face to face stacking interactions are referred to simply as stacking as it is the most commonly seen stacking interaction.

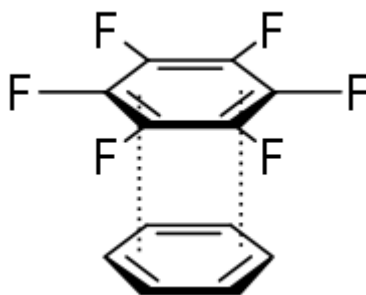


Figure 1.15. An example of an aromatic stacking interaction.

The last two interactions of this type that are commonly seen are anion- π and cation- π interactions. These involve the interaction of an anion and a cation with a π system, respectively.

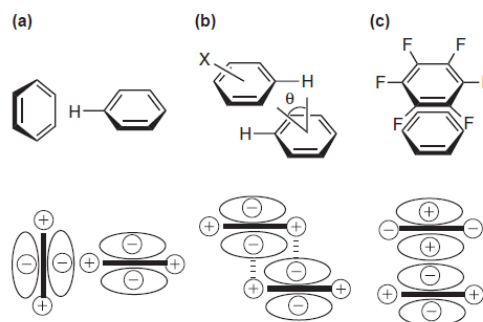


Figure 1.16. Stacking interactions (a) C-H- π stacking interaction (b) edge to edge stacking interaction and (c) face to face stacking interaction [22].

There are three main possibilities for the stacking interactions presented in this thesis. The first is edge to face stacking interaction. This interaction has one molecule lying perpendicular to another and the interaction forming between the molecules (Figure 1.16 (a)) and can be considered a CH- π interaction. The offset stacking interaction orientation is commonly found in proteins and is the geometry of DNA base stacking (Figure 1.16 (b)). In this geometry, more surface area is buried, and the van der Waals and hydrophobic interactions are increased. This orientation is seen more commonly when the electron density on the face of one or both rings is reduced. The final possible geometry is the face to face stacked orientation (Figure 1.16 (c)). This is commonly observed with donor-acceptor pairs and compounds that have opposite quadrupole moments, such that the interaction between the faces of the rings is attractive.

1.8.2 Van der Waals

In its simplest definition, van der Waals interactions are the sum of the attractive force or repulsive force between molecules other than those due to covalent bonds or to the electrostatic interaction of ions with one another or with neutral molecules. Van der Waals interactions include the force between a permanent dipole and an induced dipole and the force between two instantaneously induced dipoles. Van der Waals forces are relatively weak when compared to covalent and ionic bonding. The interactions are different from covalent and ionic bonding as van der Waals interactions are caused by correlations in the fluctuating polarisations of nearby particles.

1.9 Nucleobases

The most targeted aspect in research for new treatments of cancers and viruses is to target the bases that are present in DNA. The idea behind targeting these materials is to use a foreign element attached to the natural base to try and interrupt the base pairing that occurs in DNA or RNA.

1.9.1 Cytosine

The cytosine molecule is one of the naturally occurring base molecules found in DNA and RNA (Figure 1.17). This is found in human DNA and RNA. The crystal structure of cytosine monohydrate was first published by Jeffrey in 1963 [59] and this has since been re-determined four times [60,61,62,63]. Even with the presence of a water molecule, two different base-pairing motifs are still exhibited in the structure. These form into homo-base paired chains extending throughout the structure with the water molecule acting as both a link between the layers and between chains in the same layer (Figure 1.18).

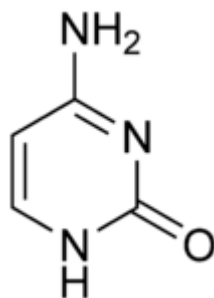


Figure 1.17. The molecular structure of cytosine.

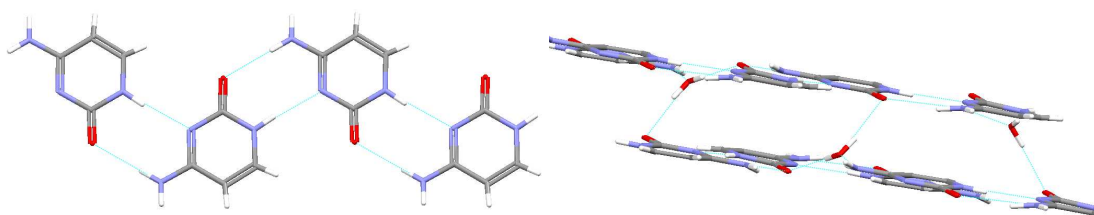


Figure 1.18. The structure of cytosine monohydrate showing the extended homo-base paired chain (left) and the water linking the layers (right) [64].

The anhydrous crystal structure of cytosine was reported in 1964 by Barker [64]. The same cytosine chain found in the monohydrate structure was also found in the anhydrous structure. In the anhydrous structure, the cytosine molecules are directly linked to the different cytosine chains in the structure through hydrogen bonds. However, unlike the

monohydrate structure which is planar with regard to the cytosine molecules, in the anhydrous structure the adjacent chains are changed in orientation so a connection can be made in the absence of the water molecule. This results in spirals that link several chains together as seen in Figure 1.19.

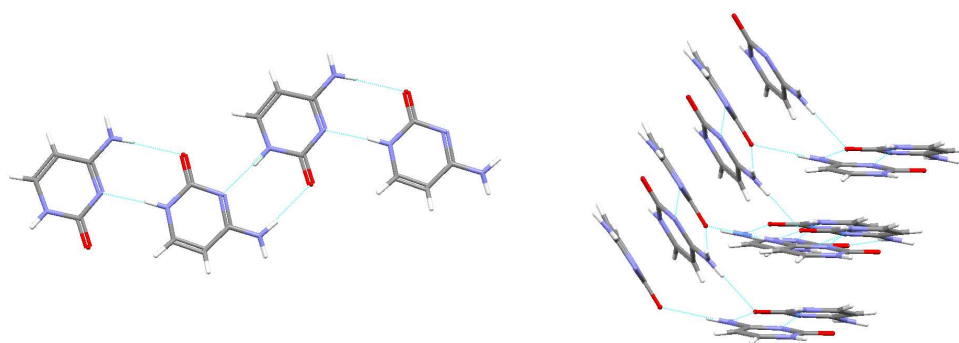


Figure 1.19. The structure of anhydrous cytosine showing the extended homo-base paired chain (left) and the chains changing orientation to link separate layers (right).

The cytosine molecule has also been widely used [65-76] as a co-molecule in the formation of molecular complexes due to the hydrogen bonding potential it possesses. One factor that has consistently affected the structure of cytosine molecular complexes is proton transfer from one molecule (for example an acid molecule) to the normally unprotonated nitrogen atom in the cytosine ring.

In 2008 Thomas *et al* [76] published a paper into the study of hydrogen bonding in proton-transfer complexes of cytosine with trimesic acid and pyromellitic acid. A molecular recognition process facilitated by the proton-transfer was exhibited in all of the structures. Both of the molecular complexes contained only charged cytosine molecules. A hydrogen bonded ring was formed between the area that had lost a hydrogen from the acid and the area of the cytosine molecule that had gained the hydrogen. The two structures contained in the paper both exhibit this bonding motif, shown in Figure 1.20.

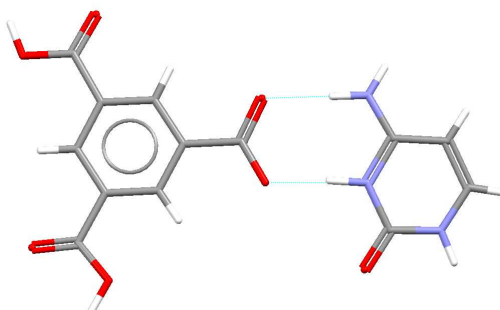


Figure 1.20. Asymmetric unit of the crystal structure of the molecular complex of cytosine and trimesic acid reported by Thomas, showing the hydrogen bonded ring bonding motif [76].

However, some structures that involve a cytosine molecule can have mixed protonation states as well. An example of this was published in 2005 by Perumalla *et al* [77] and is a

1:1 molecular complex of cytosine and benzoic acid. In this structure one of the cytosine molecules is protonated and one is neutral. One of the natural bonding motifs for bases is a three hydrogen bond base-pairing motif seen between a cytosine molecule and a guanine molecule. With the transfer of the hydrogen to one but not both of the cytosine molecules in the structure the pseudo-Watson-Crick hydrogen bonding motif [77] becomes possible as a third hydrogen bond between the newly protonated nitrogen on a charged cytosine ring and the normally unprotonated nitrogen of a neutral cytosine can now form.

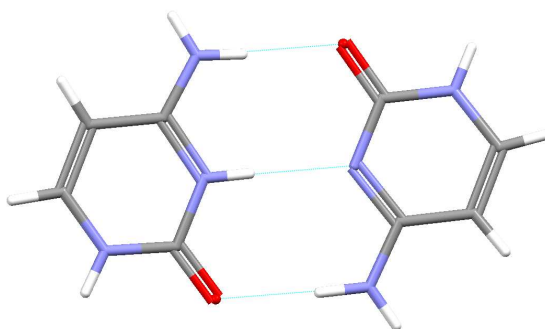


Figure 1.21. The pseudo-Watson-Crick bonding motif exhibited by the cytosinium-hydrogen maleate-cytosine molecular complex [71].

This bonding motif was later seen in the molecular complex of cytosinium-hydrogen maleate-cytosine which was published in 2009 by Benali-Cherif [71]. Once again, this structure contains a mixed protonation state for the cytosine molecule and as a result forms the pseudo-Watson-Crick unit.

1.9.2 5-Fluorocytosine

5-Fluorocytosine is a synthetic antimycotic drug (Figure 1.22) [77]. It is structurally very similar to the cytosine molecule with a fluorine atom attached to the cytosine ring instead of a hydrogen atom to one of the ring carbon atoms in the 5 position. It is widely available in an oral form and in some countries as an injectable form. This molecule was first synthesised in 1957 [77] but its antifungal properties were not truly discovered until 1964 and it was finally used on humans in 1968 [77].

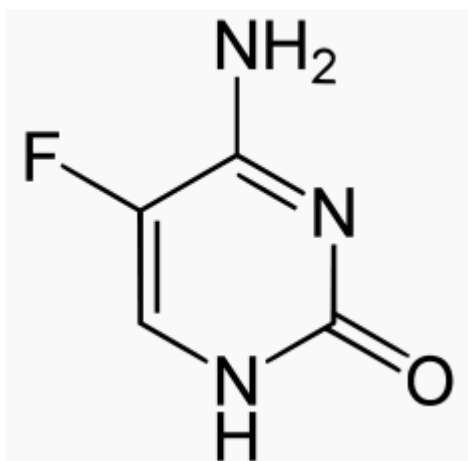


Figure 1.22. The molecular structure of 5-fluorocytosine.

The structure of 5-fluorocytosine monohydrate was first characterised by X-ray diffraction in 1982 by Louis *et al* [79]. The structure is directly comparable to the cytosine monohydrate structure and the same homo-base paired chains of 5-fluorocytosine are formed (Figure 1.23). The role that the water molecules play in both the cytosine and 5-fluorocytosine structures is similar in connecting the layers together, however the two structures are not isostructural.

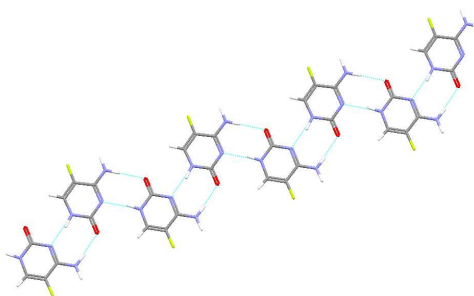


Figure 1.23. Extended homo-base paired chain found in 5-fluorocytosine monohydrate [79].

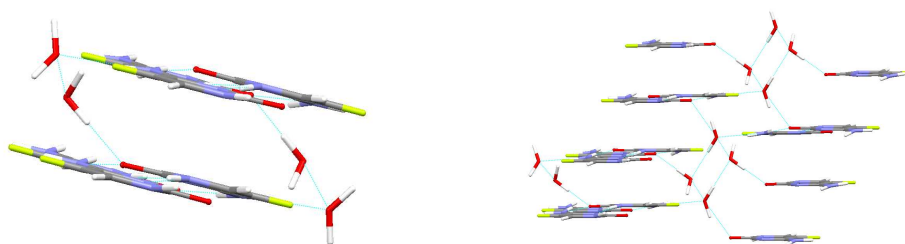


Figure 1.24. Water molecules linking separate layers found in 5-fluorocytosine monohydrate [79].

As more investigation was undertaken on the crystallisation of 5-fluorocytosine an anhydrous structure of 5-fluorocytosine was reported in 2006 by Hulme *et al* and a number of other solvated structures were also determined [80]. The crystal structure of anhydrous 5-fluorocytosine also shows the extended chains held together by base pairing (Figure 1.25). The anhydrous 5-fluorocytosine crystal structure adopts a much more layered

structure than the equivalent cytosine structure due to the presence of weak C-H...F hydrogen bonds between chains.

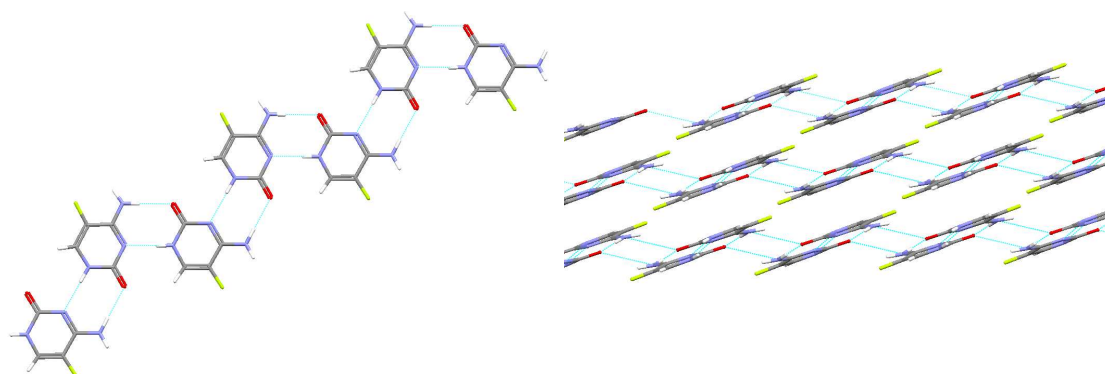


Figure 1.25. Homo-base paired chain (left) and the stacking of the bases (right) found in the anhydrous form of 5-fluorocytosine [80].

Molecular complexes of 5-fluorocytosine with various co-molecules have also been reported. One example is the molecular complex of 5-fluorocytosine and salicylic acid. This was first published in 2001 by Prabakaran *et al* [81] and is more precisely named 5-fluorocytosinium salicylate.

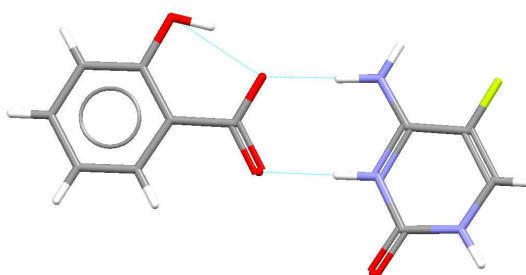


Figure 1.26. Hydrogen bonded ring found in the molecular complex of 5-fluorocytosine and salicylic acid [81].

An important factor in this crystal structure is the transfer of a hydrogen atom from the salicylic acid to the 5-fluorocytosine molecule. This means that both the molecules in the structure are charged and these combine to form a dimer involving the area of the salicylic acid that has lost the hydrogen atom and the area of the 5-fluorocytosine that has gained the hydrogen atom. This forms a hydrogen bonded ring (Figure 1.26). Chains are formed via single hydrogen bonds linking separate dimers together (Figure 1.27).

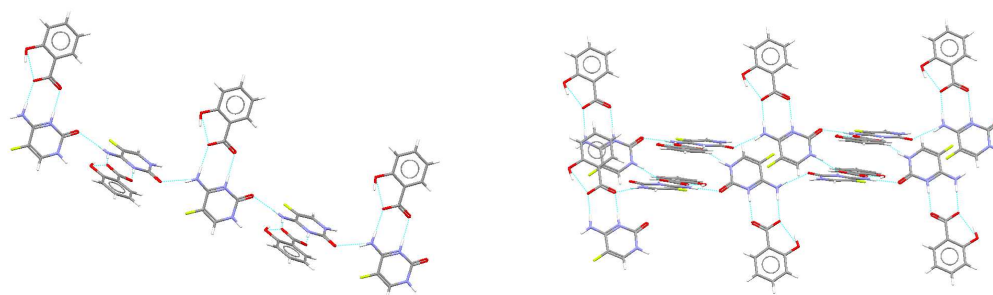


Figure 1.27. Single chain (left) and two linked chains (right) found in the molecular complex of 5-fluorocytosine and salicylic acid [81].

1.9.3 Thymine

In 1961 Gerdil reported the monohydrate structure of thymine [82]. The thymine molecules combine to form an extended chain through homo-base paired units as shown in Figure 1.28. This chain is very planar and contains water molecules tied to the chain on either side via hydrogen bonds as seen in Figure 1.29. These water molecules play an important role in linking separate chains of thymine molecules in separate layers. Adjacent chains do not lie parallel to one another.

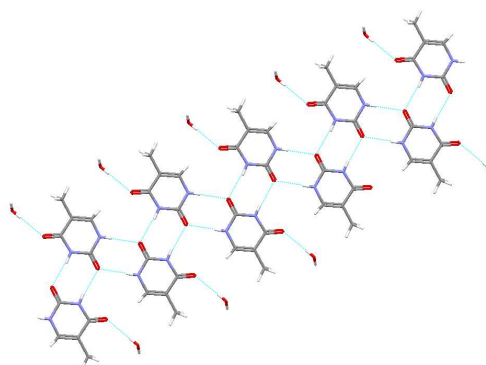


Figure 1.28. Homo-base paired chain found in thymine monohydrate [82].

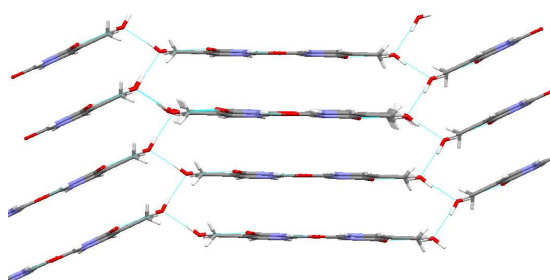


Figure 1.29. Water molecules linking separate layers in thymine monohydrate [82].

The co-crystallisation of thymine has not been widely investigated but molecular complexes with 2-(acrylamido)-6-(methyldamido)pyridine and p-benzoquinone have been reported [83,84].

1.9.4 Adenine

The structure of adenine trihydrate has been reported in 1987 by S.M. Tret'yak [85]. No anhydrous forms have been reported to date. Adenine trihydrate does not show any homo-base pairing, a likely effect of the increased proportion of water molecules contained within the complex. The water molecules encircle all of the adenine molecules (Figure 1.30) and the only direct interactions between adenine molecules are base-stacking between layers (Figure 1.31).

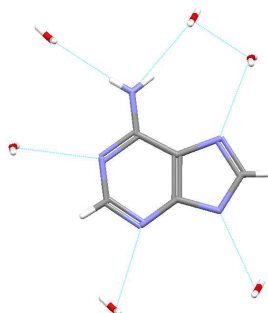


Figure 1.30. The hydrogen bonding involving the adenine molecule in the trihydrate structure. No homo-base pairing is found [85].

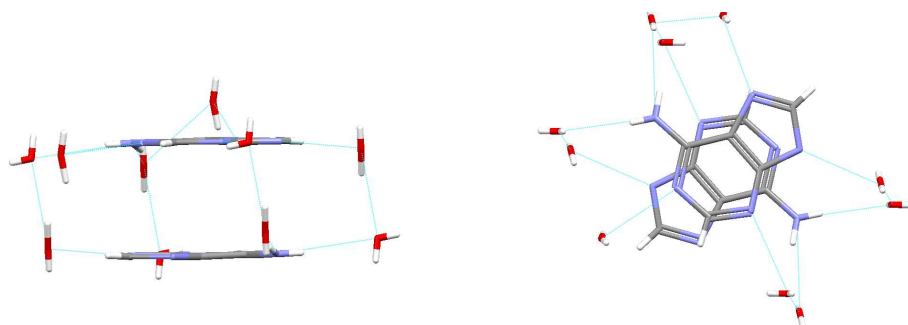


Figure 1.31. Water molecules linking separate layers in the molecular structure of adenine hydrate [85].

Adenine has been the target of several co-crystallisation attempts due to the excellent hydrogen bonding opportunities that are available when this molecule is used. In 2007 Sridhar [86] reported a structure that contained the adenine molecule co-crystallised with phthalic acid. As seen in some of the cytosine structures, hydrogen transfer occurs from one of the carboxylic acid groups of the phthalic acid to the adenine molecule. In the structure, the crystallisation ratio is 1:1 but there are two of each type in the asymmetric unit; all of the adenine molecules have gained a hydrogen atom whilst one phthalic acid molecule has been deprotonated and the other has remained neutral.

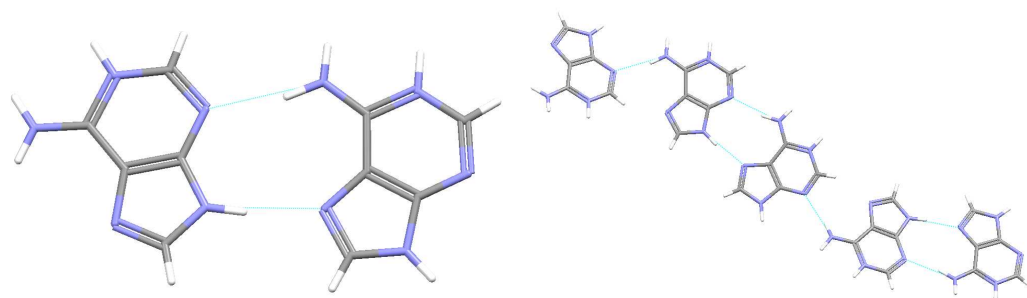


Figure 1.32. Homo-base paired chain in molecular complex of adenine with phthalic acid [86].

In this molecular complex, unlike what was seen in the adenine trihydrate structure, there is homo-base pairing (Figure 1.32, left). These base pairs are linked together via single hydrogen bonds to create an extended chain made up of the adenine molecules (Figure 1.30, right).

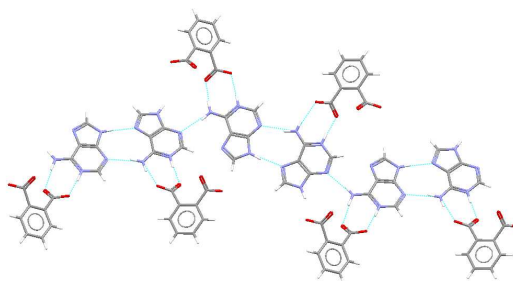


Figure 1.33. Phthalate molecules tied to the homo-base paired chain in molecular structure of adenine with phthalic acid [86].

On either side of the chain there is a phthalate molecule tied to the chain via a hydrogen bonded ring motif. This is made up of two hydrogen bonds between two molecules and this occurs on both sides of the chain always involving the same region of the adenine molecule and the caboxylate group of the phthalate. The twisting of the carboxylate groups of the phthalate molecule allows the phthalate molecules to form hydrogen bonded rings with two separate layers using the two carboxylate groups as shown in Figure 1.34. There are also base stacking interactions between the layers.

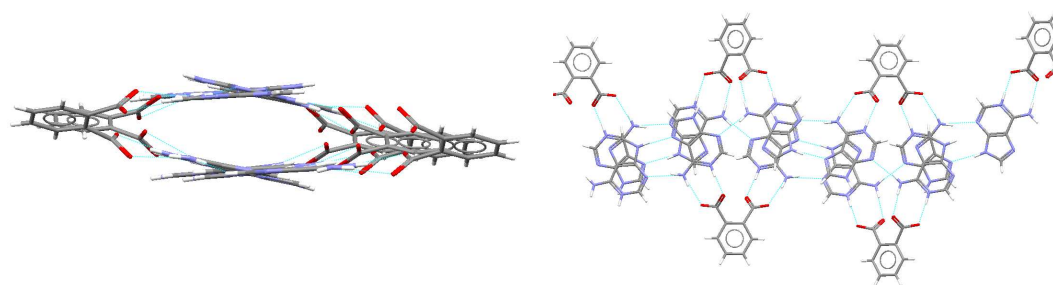


Figure 1.34. Phthalate molecules linking separate homo-base paired chains in molecular complex of adenine with phthalic acid [86].

It was not until 2008 that the pure structure of adenine was reported. This structure was published after work carried out by Mahapatra et al [87]. This structure contains hydrogen bonded base pairs with the two unique base pairs (Figure 1.35) being made up of two hydrogen bonds each. One of the base pairs involves the equivalent face of two bases whilst the second base pair involves two distinctly different faces of the bases. These base pairing motifs combine to form sheets of adenine which are approximately planar in nature (Figure 1.35, right).

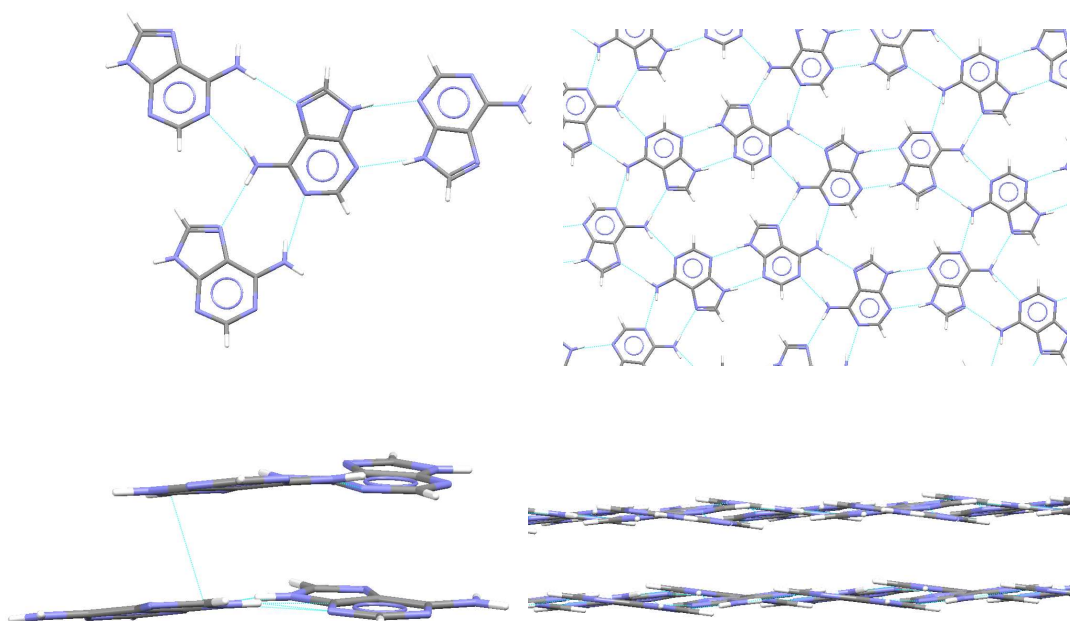


Figure 1.35. The base pairing of adenine molecules into layers [87].

The planes are linked together via base stacking interactions. It has been commonly seen in the other pure base structures that form sheets that the main connection between separate layers are base stacking interactions and that these stacking interactions are staggered between the layers.

1.9.5 Guanine

The crystal structure of guanine monohydrate was reported by Thewalt in 1971 [88]. There is no known anhydrous structure. Unlike the structure of adenine trihydrate, guanine monohydrate does contain homo-base pairing, a likely result of the decreased relative amount of water (Figure 1.36).

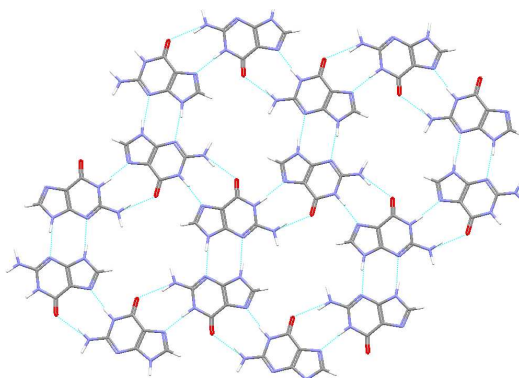


Figure 1.36. Guanine molecules homo-base pairing to form sheets in guanine monohydrate [88].

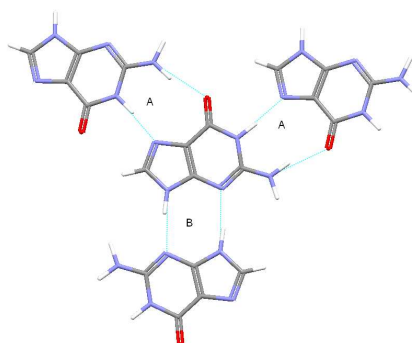


Figure 1.37. Two different homo-base pairing motifs in guanine monohydrate [88].

There are two unique base pairing motifs present in this structure. These are depicted in Figure 1.37 (A and B). It is a combination of the two unique bonding motifs that help to form the extended planes as shown in Figure 1.38. In this Figure, it can be seen that there are small “voids” present in the structure. However, these are not true voids and are actually filled with the water molecules which link the layers together. The water molecules form hydrogen bonded chains that runs perpendicular to the planes generated by guanine molecules (Figure 1.38).

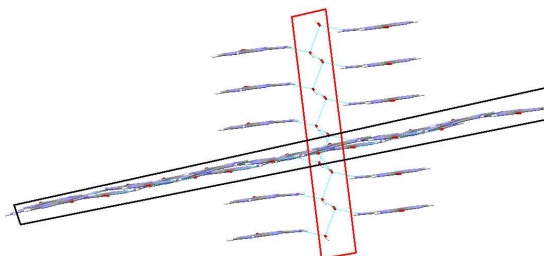


Figure 1.38 Water molecules forming a chain that links different layers of the guanine sheets in the molecular complex of guanine [88].

Guanine is one of the lesser targeted molecules for co-crystallisation of the bases due to its solubility issues. This has meant that a cytosine and guanine molecular complex has not been reported to-date. As a result the use of guanine as a co-molecule is limited with only a

few structures entered into the CSD. One of these structures was published in 1975 by Bugg, the molecular complex guanine picrate monohydrate [89].

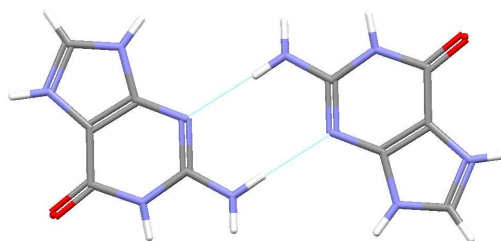


Figure 1.39. Homo-base pair present in the molecular complex of guanine picrate monohydrate [89].

Even with the addition of a co-molecule and water molecule there is still homo-base pairing in this structure (Figure 1.39). This base pair is made up of two hydrogen bonds but is not equivalent to the base-pairing motifs seen in the guanine monohydrate structure. The formation of an extended guanine chain is terminated by the combined effect of the picrate and the water molecules (Figure 1.40).

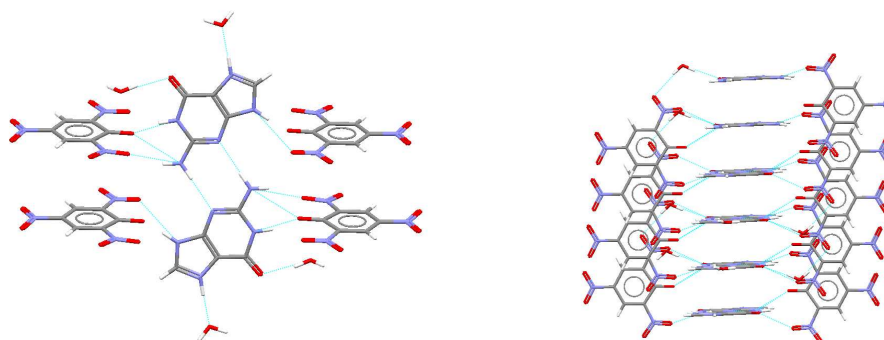


Figure 1.40. Picrate molecules combining with water molecules to terminate the guanine homo-base paired chain in the molecular complex of guanine picrate monohydrate [89].

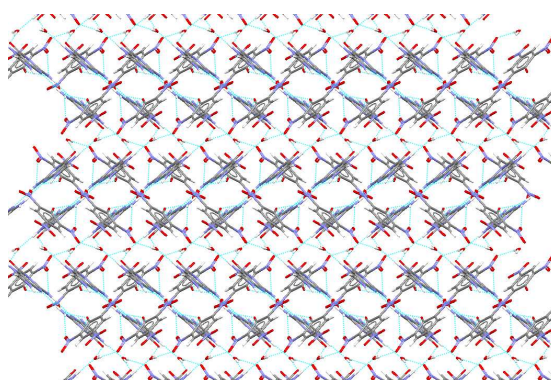


Figure 1.41. Crystal packing in the molecular complex of guanine picrate monohydrate [89].

1.3.6 Uracil

There are four base molecules that are present in DNA. In RNA, thymine is replaced by uracil. This base molecule is possibly the material with the most research conducted on it. In 1967 Stewart published the crystal structure of uracil [90].

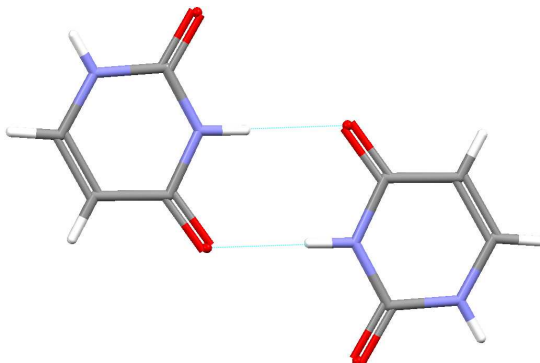


Figure 1.42. Homo base pair in the crystal structure of uracil [90].

There is homo-base pairing in this structure made up of two hydrogen bonds between two equivalent uracil molecules (Figure 1.42). These base pairs are linked to other base pairs via single hydrogen bonds as seen in Figure 1.43 (left). These chains are linked to other chains to create sheets of uracil molecules that are very planar and can clearly be seen in Figure 1.43.

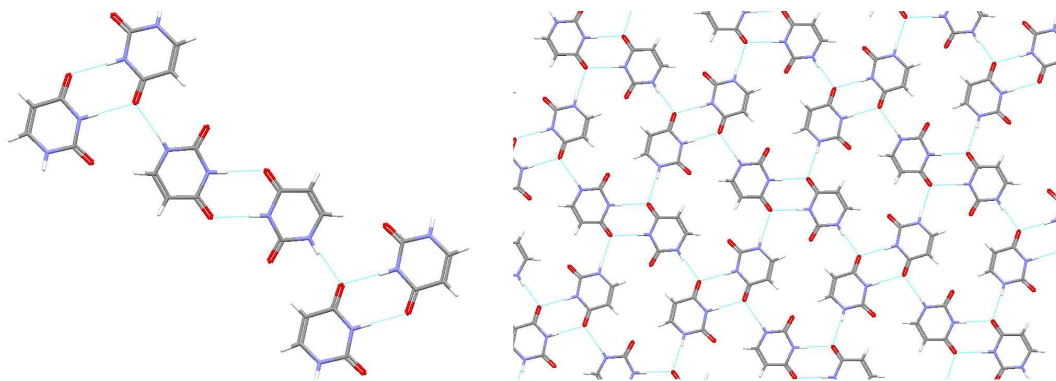


Figure 1.43. Homo base chain (left) and uracil sheets (right) in the crystal structure of uracil [90].

The sheet of uracil molecules are held together via base-stacking interactions (Figure 1.44).

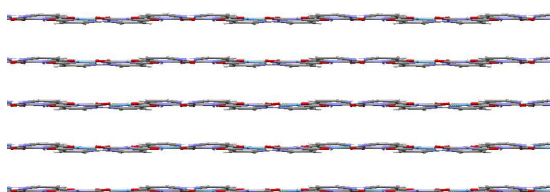


Figure 1.44. Base stacking interactions between layers in the crystal structure of uracil [90].

Uracil has also been studied in co-crystallisations; uracil differs from cytosine as it is not possible for uracil to gain a hydrogen atom through proton transfer. One of the most recent structures reported regarding uracil was published in 2007 by Thomas *et al* [91]. This structure was between melamine and uracil and is produced in a 1:1 ratio. In this structure there is a bonding motif that resembles a Watson-Crick hydrogen bonded unit between a uracil molecule and a melamine molecule. This bonding motif is made up of three hydrogen bonds and is depicted in Figure 1.45.

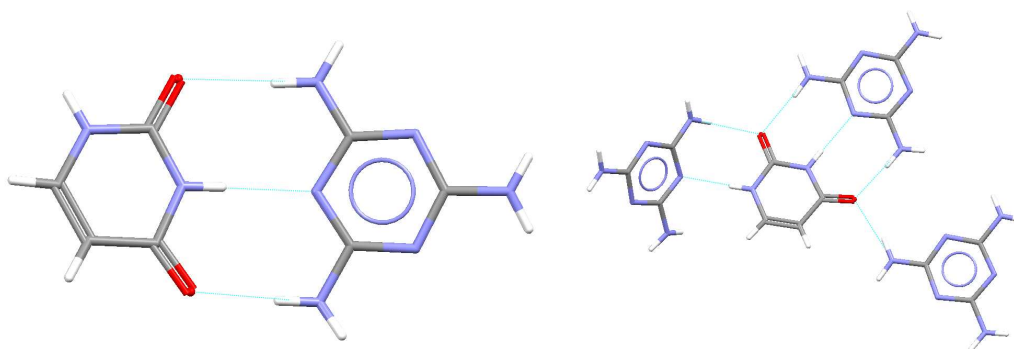


Figure 1.45. Hydrogen bonding motif (left) and all hydrogen bonding involving uracil molecule (right) in the molecular complex of uracil and melamine [91].

There is no homo base pairing in this structure at all as the uracil molecule hydrogen bonds solely with melamine molecules as depicted in Figure 1.45 (right). However, an extended chain is formed with a twisted melamine melamine hydrogen bonding motif linking separate layers in the structure. This provides strong three dimensional links between the layers in addition to base stacking interactions.

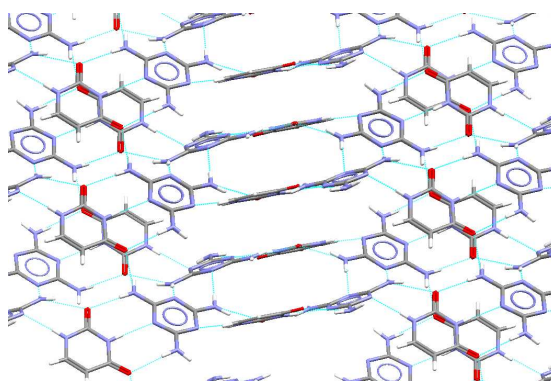


Figure 1.46. Three-dimensional packing in the molecular complex of uracil and melamine [91].

1.9.7 5-Fluorouracil

5-Fluorouracil is the fluorinated derivative of uracil. This molecule has had a wide variety of research conducted in particular with regards to the anti-viral and anti-cancer properties that it possesses [92,93]. In 1973, Fallon first published the crystal structure of 5-fluorouracil [94]. In 2005, a polymorph of this structure was published by Hulme *et al*

[95]. Both polymorphs show assembly of 5-fluorouracil molecules into chains, however, the construction of these chains differs between the two. The first polymorph has more in common with the crystal structure of uracil in that it forms sheets of 5-fluorouracil molecules where base paired units are connected by single hydrogen bonds into layers (Figure 1.47).

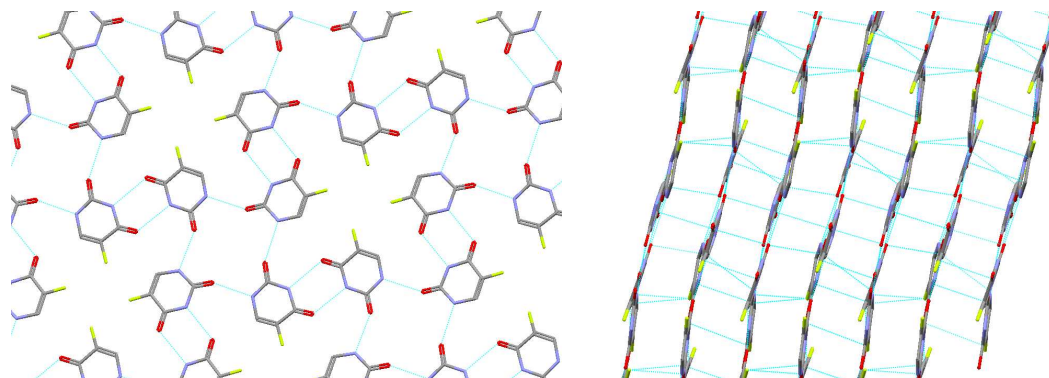


Figure 1.47. Sheets of 5-fluorouracil (left) packed into layers (right) in the crystal structure of 5-fluorouracil form I [94].

The second polymorph shows chains of extended homo-base-pairing (Figure 1.48) [95]. The fluorine atoms form interactions between chains and this is key to the formation of the layers (Figure 1.49).

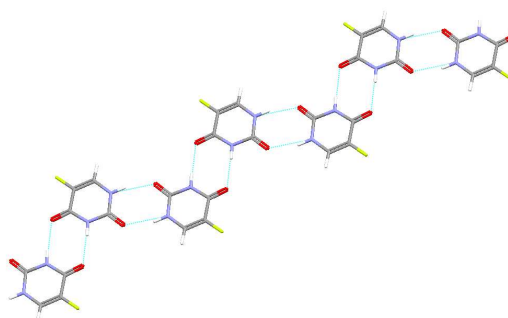


Figure 1.48. Chains of 5-fluorouracil in the structure of 5-fluorouracil form II [95].

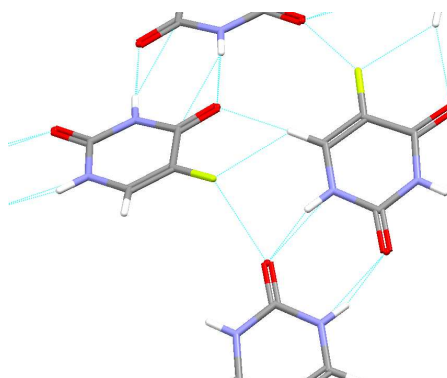


Figure 1.49. Fluorine interactions in the structure of 5-fluorouracil form II [95].

5-fluorouracil has also been co-crystallised with a number of co-molecules. In 1969 a molecular complex of 5-fluorouracil cytosine monohydrate was reported by Voet *et al* [67].

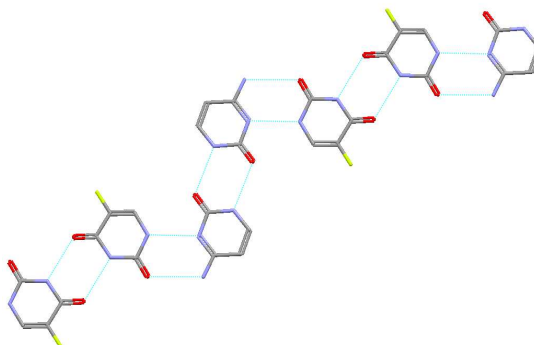


Figure 1.50. Base paired chain in the molecular complex of cytosine and 5-fluorouracil monohydrate [67].

A chain of alternating homo-base-paired 5-fluorouracil and cytosine units is formed where the two units are connected by hetero-base-pairing between cytosine and 5-fluorouracil (Figure 1.50).

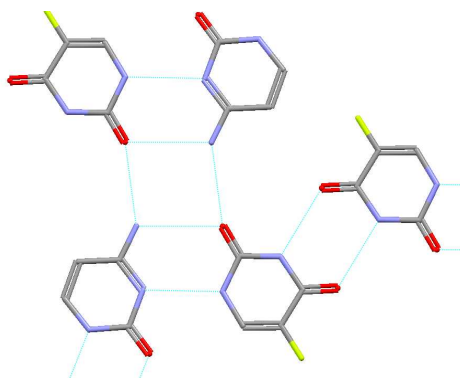


Figure 1.51. Hydrogen bonding linking the chains in the molecular complex of cytosine and 5-fluorouracil monohydrate [67].

Separate chains are linked together via single hydrogen bonds between a cytosine molecule and a 5-fluorouracil molecule. This results in the formation of a hydrogen bonded diamond (Figure 1.51). The layers are linked via base attacking interactions.

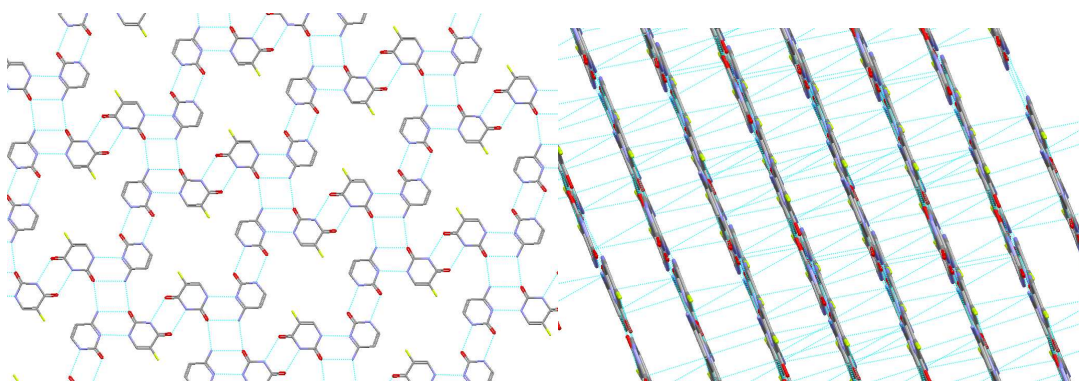


Figure 1.52. Sheets of 5-fluorouracil and cytosine (left) and base stacking (right) in the molecular complex of cytosine and 5-fluorouracil monohydrate [67].

1.10. Classification of Base Pairs

Base paired units can be classified by the type of hydrogen bonding motif displayed and also by the relative orientations of the two bases comprising the pair. The classifications of the hydrogen bonded units referred to throughout this work are defined here. A detailed discussion of the hydrogen bonding motifs adopted by cytosine in molecular complex formation is given in Chapter 8.

1.10.1 Pseudo-Watson-Crick Base Pair.

One of the common bonding motifs seen in the CSD is the pseudo-Watson-Crick bonding motif [97]. This bonding motif is the equivalent of the cytosine and guanine base pair. It consists of three hydrogen bonds with the most common arrangement being two N-H...O hydrogen bonds) surrounding an N-H...N hydrogen bond (Figure 1.53).

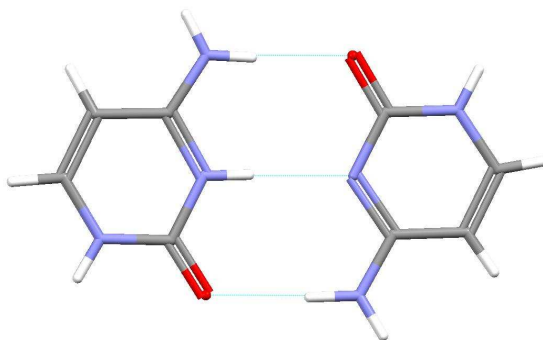


Figure 1.53.An example of a pseudo-Watson-Crick bonding motif [101].

This bonding motif has been regularly seen between two cytosine molecules in the CSD, however, this can only occur between two cytosine molecules if the structure contains a charged cytosine (gained a hydrogen to the normally non-protonated heteroatom in the ring) and a neutral cytosine. This bonding motif is restricted by the presence of a neutral cytosine molecule and a charged cytosine molecule being present in the structure for this structure to happen. This bonding motif has been widely seen before with much work done on it [98].

1.10.2.Pseudo-Hoogsteen Base Pairing

These three bonding motifs are equivalent to the bonding seen between adenine and thymine with only two hydrogen bonds making up the base pair.

1.10.2.1 Type A

In a type A bonding motif the bonding between the two molecules involved consists of N-H...N hydrogen bonds (Figure 1.54).

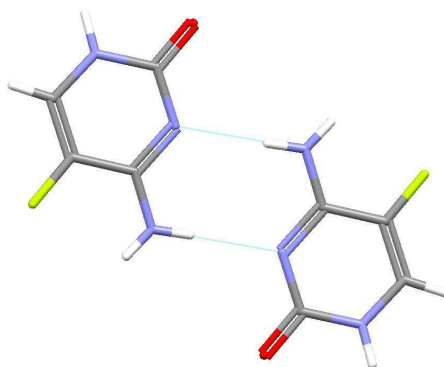


Figure 1.54. An example of a typical type A bonding motif [99].

This bonding motif is not as common in the CSD as the pseudo-Watson-Crick. This may be due to the fact that this tends to form in neutral complexes and a large number of structures in the CSD contain hydrogen transfer. There is limited work that has been done on structures that contain this bonding motif [99].

1.10.2.2 Type B

The most common out of all the hydrogen bonding motifs is the type B bonding motif. This bonding motif also consists of two hydrogen bonds between the two molecules, however, in this case is is formed by two O-H...O hydrogen bonds (Figure 1.55).

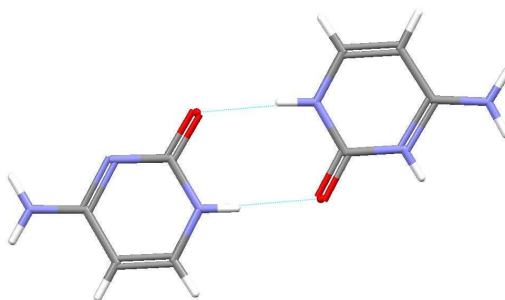


Figure 1.55. An example of a typical type B bonding motif [100]

This bonding motif can combine with any of the other bonding motifs (not possible with type C) to create extended chains. This bonding motif often acts as a linker between other base pair motifs, most predominantly the pseudo-Watson-Crick [100].

1.10.2.3. Type C

This bonding motif is a cross between type A and type B using one of the faces from each to make up a bonding motif. This bonding motif is not very common in the CSD but appears in the structure of cytosine monohydrate [99].

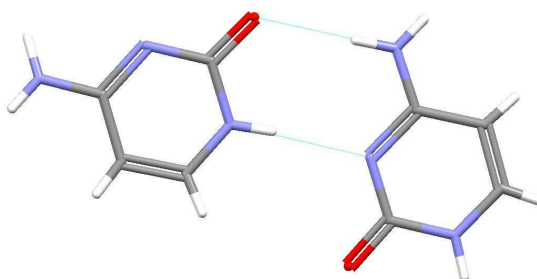


Figure 1.56. An example of an uncommon type C bonding motif [99].

This bonding motif uses one face commonly used in the type A bonding motif and one face used in the type B bonding motif. This bonding motif can be defined as being made up of two hydrogen bonds, one N-H...N and one O-H...O hydrogen bond.

1.10.3. Hydrogen Bonded Hetero-pseudo-dimer

When a ‘traditional’ base-pairing motif is not formed then a bonding motif made up of one base and one non-base will predominantly form. This bonding motif is formed commonly when there is hydrogen transfer. This bonding motif will normally involve the part of the molecule that has lost the hydrogen (the acid) and the part of the molecule that has gained the hydrogen in the base. There is no restriction on the type of hydrogen bonds making up this motif, however, the most commonly seen bonding motif has two O...H-N hydrogen bonds.

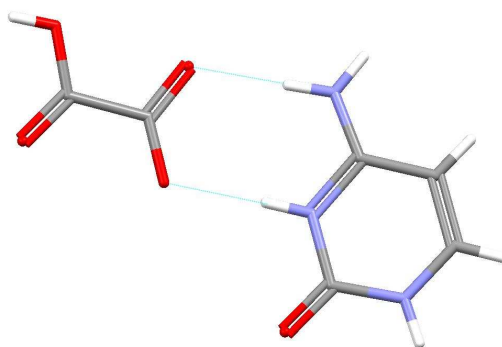


Figure 1.57. An example of a typical hydrogen bonded heterodimer [100]

The most common occurrence is the transfer of a hydrogen from a carboxylic acid group to the normally unprotonated nitrogen of the cytosine. This allows for the formation of a hydrogen bonded ring involving the carboxylate group and the amine group and the newly protonated nitrogen of the cytosine ring [100].

1.10.4 No Primary Base Pairing Motif

Some of these hydrogen bonding motifs are more common than others but the majority of the structures in the CSD contain at least one of the primary base pair hydrogen bonding motifs. However, there are some structures where no base pairing motif is found and only single hydrogen bonds involving the base occur [101].

1.11 Conformational parameters

Much work has been carried out to analyse the different effects that the packing of a crystal structure and the basic build up can have on a base-paired motif. In 1987 Wilson [103] undertook work that analysed hits in the CSD. This work was done to see the range of effects that can be found in crystal structures containing nucleosides and nucleotides. The main parts of a base-pair that can be analysed are the hydrogen bond parameters, propeller twist, buckle and the C1'-C1' separation.

The propeller twist can be described as the angle between the planes of the two bases that are involved in the base-pairing motif when viewed along the axis that is connecting them. It can be seen as positive when the rear base has to be rotated clockwise for the bases to be flattened out. In helical structures the propeller twist is always positive but in nucleoside and nucleotide structures the sign will usually be seen an equal number of times between negative and positive. The buckle of a base-pair can only be seen once the propeller twist has been flattened. This is the dihedral angle between the two planes along the short axis. The C1'-C1' is the separation of the C1' of one base to the C1' on the other base material. Finally, the hydrogen bonded base-pair can be of several types and depends on the base

materials that are involved (Watson-Crick, Hoogsten etc.) and the angles and lengths can be analysed.

The work carried out here showed that the variation of the propeller twist did not depend on the other parameters stated. This seemed a remarkable result mainly that the strength of the hydrogen bonds in the base-pair has no effect on the magnitude of the propeller twist. In the results it was seen that the propeller twist could vary from 0.8° to 49.1° . These results agreed with previous work carried out by Wilson and Tollin [102] that nucleoside base-pairs have a natural propeller twist that can be large even in the absence of the standard DNA structure and in varying crystal structure packing. However, the average value for the propeller twist in this study was found to be 14.7° . The purine bases have a larger twist (on average 15.7°) when compared to those seen in pyrimidine bases (12.1°). The normal range of a propeller twist can be seen to be from 10° to 20° . The largest propeller twist is from adenine structures with the smallest twist relating to cytosine structures. However, there is no similar trend seen in the buckle of C1'-C1' separation.

There was further work conducted in 1990 by Wilson [104] on the effect of exocyclic substituents on the base-pair propeller twist. This work reinforced the idea that the base pairs are normally non-planar units. With the analysis it has been seen the naturally adopted conformation could possibly be the non-planar axes with the twisting once again occurring about their long axis. This questions the common assumption that isolated base-pairs are normally co-planar and that the twisting is a reaction to the base stacking interactions that are present in the structure. This work also put forward the idea that the twisting was smaller in structures that contain only bases may be related to the planarity of the group stacking present when compared to the twist seen in oligonucleotides.

In 1987 Hobza and Sandorfy [105] conducted nonempirical calculations on all 29 possible base pairs. The work carried out here showed that there was good agreement between the experimental gas-phase enthalpies and theoretical enthalpies at 300K has been obtained for GC, CC, AT and TT pairs. The GC pair, with three hydrogen bonds present between the two molecules is the most stable. However, as seen the stability is also related to the base stacking interactions and not reliant on the presence of the three hydrogen bonds. The GG pair, even with only 2 hydrogen bonds, is comparably stable. The molecules that make up the pair, even if the bonding motif is the same as another, will have a great effect on the stability. This means that neither the number nor the linearity of the bonding motif is primarily responsible for the stability of the base pair. This work also showed that the stability cannot only take into account the atoms making up the bonding motif but the whole molecule must be analysed.

This work also concluded that the hetero base pairs are more stable than the homo base pairs with complementary base pairs and that homo base pairs are more stable than hetero base pairs when there is non-complementary base pairing.

Chapter 2 – Techniques

2.1 Introduction to Crystallography

Crystallography is the science of crystals. It has to do with the external shape, the internal structure, the growth and the physical properties of crystals. When crystallography was first introduced it was purely descriptive and was considered a branch of mineralogy. It was later discovered that the crystalline state is not confined to mineral salts alone and it is actually a very common state of matter. It was not however, until the middle of the nineteenth century that crystallography had become a science in its own right.

For a long time it was thought that the external appearance of crystals reflected some regular internal ordering of matter. The first references to this can be found in work by Johannes Kepler, Robert Hooke and Christian Huyghens [106-108]. Huyghens' work involving the birefringence of calcite suggested that its optical properties could be explained by the rules governing the internal arrangement within the crystal [108].

In 1669, the first quantitative law of crystallography (the law of constant angles) was determined by Nils Steensen from the measurement of the angles between the faces of a quartz crystal [109]. It was not until 1772 that the law was formally expressed by Jean Baptiste Rome de l'Isle in his 'Essai de Cristallographie' [110].

The second law, the law of rational indices or simple truncations, was first stated in 1774 by the Abbe Rene-Just Haüy [111]. It had been noticed that when a crystal of calcite was cleaved, the pieces that were obtained had shapes identical to that of the original piece. The assumption was made that the crystals were made up of identical parallelepipeds which he called 'molecules integrantes'. This led to the understanding that the position in space of each face of a crystal could be described by three numbers. These ideas were later refined by W.H. Miller, who introduced the method of analytical geometry into crystallography and who proposed the notation system, "Miller indices", which is still in use [112].

The contribution made by Auguste Bravais to crystallography is of massive importance to this area. In 1849 in his work 'The Lattice Structure of Crystals', he stated the following postulate:

"Given any point **P** in a crystal, there exists an infinite number of discrete points, unlimited in the three directions of space, around each of which the arrangement of matter is the same as it is around the point **P**."

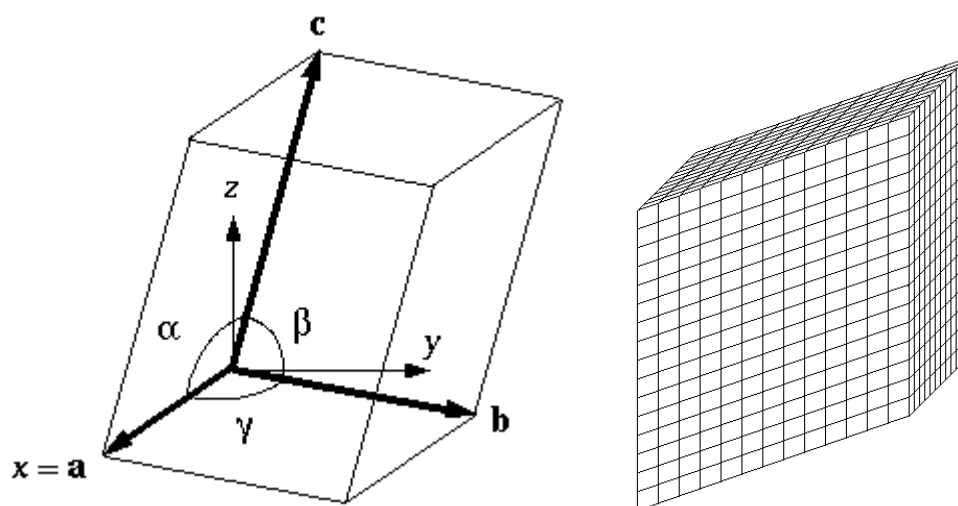


Figure 2.1. Left, the unit cell and right, the combination of many unit cells to produce the full three-dimensional crystal.

A crystal structure is a unique arrangement of atoms in a crystalline liquid or solid. A set of atoms are arranged in a particular way on a lattice exhibiting long-range order and symmetry. The crystal structure contains a small repeating unit that through symmetry, reproduces the entire crystal from this small repeating unit and that contains all of the structural information unique to the crystal. The points can be thought of as forming identical tiny boxes that are referred to as the unit cells (Figure 2.1, left). These fill the space of the lattice. The lengths of the edges of the unit cells (a , b and c) and the angles between them (α , β and γ) are called the lattice parameters. The unit cell is the smallest unit of volume that contains all of the structural and symmetry information of the crystal and that by translation can reproduce a pattern in all space (Figure 2.1, right). If the unit cell contains no symmetry then there are no restrictions on the axes and angles of the unit cell in the crystal. If there is symmetry within the unit cell, there are some restrictions on the angles and axes. There are seven crystal systems that a crystal can belong to (Figure 2.2) and these take into account all of the possible symmetry.

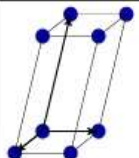
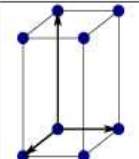
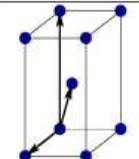
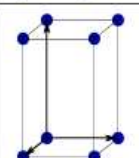
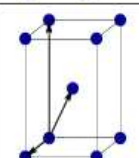
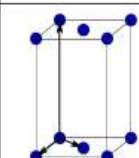
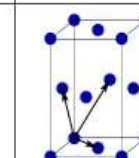
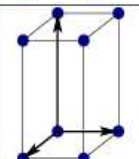
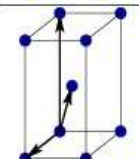
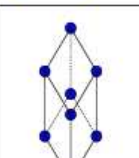
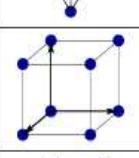
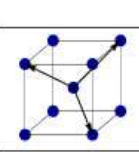
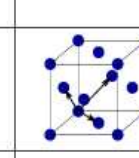
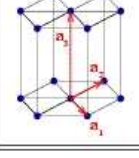
Bravais lattice	Parameters	Simple (P)	Volume centered (I)	Base centered (C)	Face centered (F)
Triclinic	$a_1 \neq a_2 \neq a_3$ $\alpha_{12} \neq \alpha_{23} \neq \alpha_{31}$				
Monoclinic	$a_1 \neq a_2 \neq a_3$ $\alpha_{23} = \alpha_{31} = 90^\circ$ $\alpha_{12} \neq 90^\circ$				
Orthorhombic	$a_1 \neq a_2 \neq a_3$ $\alpha_{12} = \alpha_{23} = \alpha_{31} = 90^\circ$				
Tetragonal	$a_1 = a_2 \neq a_3$ $\alpha_{12} = \alpha_{23} = \alpha_{31} = 90^\circ$				
Trigonal	$a_1 = a_2 = a_3$ $\alpha_{12} = \alpha_{23} = \alpha_{31} < 120^\circ$				
Cubic	$a_1 = a_2 = a_3$ $\alpha_{12} = \alpha_{23} = \alpha_{31} = 90^\circ$				
Hexagonal	$a_1 = a_2 \neq a_3$ $\alpha_{12} = 120^\circ$ $\alpha_{23} = \alpha_{31} = 90^\circ$				

Figure 2.2. The seven crystal systems and the fourteen Bravais lattices. [175]

Further to this, it is possible to define lattice types, called Bravais lattices, which are related to the internal structure of the aforementioned unit cell. A primitive unit cell has only one lattice point and all crystals can be described by a primitive lattice. However, in some cases, other non-translational symmetry can be present and it can be beneficial to define a unit cell with more than just one lattice point associated with it, centred lattices:

1. C-centred - contains an extra lattice point in the centre of the C-face (or equivalent A or B-Centred lattices);
2. I, body-centred - an extra lattice point is located in the centre of the unit cell;
3. F, face-centred - an extra lattice point is located in all faces of the unit cell.

This results in a total of 14 Bravais lattices (Figure 2.2).

There can also be additional symmetry present within the unit cell itself. The symmetry within the unit cell includes inversion, rotation, reflection or combinations of these with translation, called glide planes (reflection and translation), or screw axes (rotation and translation). The smallest motif which when acted on by a symmetry operation(s) produces the unit cell of the structure is known as the asymmetric unit. When this is combined with the lattice type, it provides the full description of the crystal. There are 32-point groups that are produced by the combination of all inversion, rotation and reflection elements together with the 14 Bravais lattices. These, combined with the elements including translation, leads to the production of 230 possible space groups, which define every possible crystal structure.

2.2 Diffraction and Structure Solution

2.2.1 X-ray Scattering

A crystalline solid is made up of an infinite number of molecules arranged in a regularly repeating pattern, generating a diffraction grating. In X-ray crystallography, X-ray waves interact with matter through the electrons, commonly referred to as Thomson scattering, contained in atoms. When the X-rays reach the electron it becomes a secondary source of electromagnetic radiation that scatters the incident radiation. The waves that are produced from the scattering from electrons can combine constructively or destructively to produce a scattering pattern called a diffraction pattern. The waves combine constructively if the waves are in phase and thus add together to produce a stronger peak and interact destructively if the waves are out of phase with one another and they subtract from each other to some degree as shown in Figure 2.3.

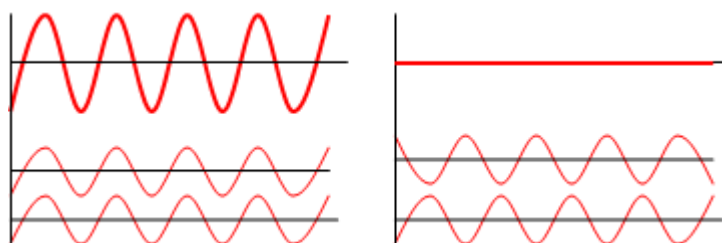


Figure 2.3. Constructive interference (left) and destructive interference (right).

The X-rays are diffracted as they pass through a diffraction grating or the lattices of a crystalline sample. These waves are then collected by a detector to produce a diffraction pattern that is seen as the output from a program used to visualise the experimental measurement (Figure 2.4).

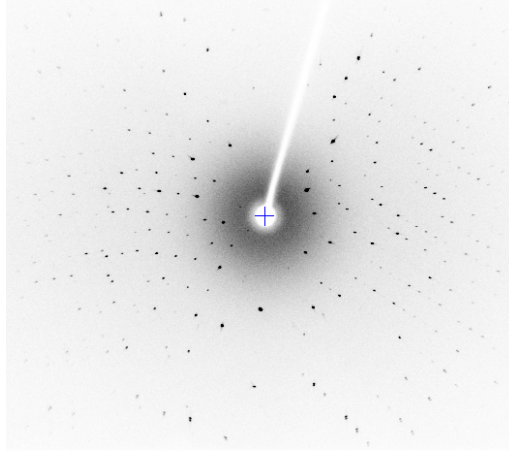


Figure 2.4. A typical diffraction pattern.

The angular spacings of the features in the diffraction pattern are inversely proportional to the dimensions of the object causing the diffraction. This means that the smaller the object causing the diffraction the wider the spacing of the peaks in the resulting diffraction pattern.

2.2.2 Diffraction

The atoms in a crystal diffract X-rays and produce a series of spots, with the positions and intensities being determined by the crystal contents and symmetry. The spots formed are from the constructive and destructive interference (Section 2.2.1) of the X-rays with each other, forming a three-dimensional reciprocal space pattern through which two-dimensional slices are measured using area detectors in a diffraction experiment. The reciprocal lattice represented in reciprocal space is governed by vectors \underline{a}^* , \underline{b}^* and \underline{c}^* (of dimension a^* , b^* and c^*):

$$\underline{a}^* = (\underline{b} \times \underline{c}) / V \quad \underline{b}^* = (\underline{c} \times \underline{a}) / V \quad \underline{c}^* = (\underline{a} \times \underline{b}) / V \quad (1)$$

where \underline{a} , \underline{b} and \underline{c} are the unit cell vectors in real space and V is the volume of the unit cell.

2.2.3 Bragg's Law

Bragg's law was derived by physicist Sir William Lawrence Bragg in 1912 [113]. This law helped to confirm the existence of particles at the atomic scale and provided a powerful new tool for studying crystals in the form of X-ray and neutron diffraction.

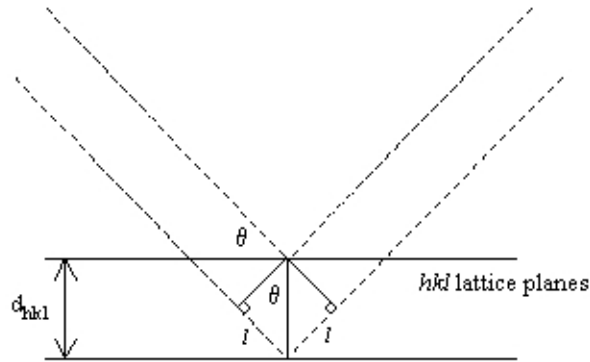


Figure 2.5. A graphical representation of Bragg's Law.

Since the quantification of Bragg's Law, which made practical crystallography a possibility, scattering techniques have been used to provide valuable information on the spatial arrangements of a wide range of materials from minerals and inorganic compounds, to molecules and macromolecules. Bragg's Law can be used to define the diffracted beam geometry:

$$n\lambda = 2d \sin \theta \quad (2)$$

where d = the distance between crystallographic planes, λ = wavelength, θ = scattering angle, n = diffraction order (Figure 2.5).

Bragg's Law results from the fact that the X-rays reflected by a stack of planes will only interfere constructively when the above equation is met. d can be calculated for all the measured stacks of planes in the crystal.

2.2.4 Miller Indices

Bragg's law is a special case of the Laue equations, which define the condition for diffraction to occur:

$$\mathbf{a} \cdot \mathbf{S} = h; \mathbf{b} \cdot \mathbf{S} = k; \mathbf{c} \cdot \mathbf{S} = l \quad (3)$$

In these equations \mathbf{a} , \mathbf{b} and \mathbf{c} are the vectors that define the unit cell of the crystal and \mathbf{S} is the vector path difference of the incident and reflected rays for the hkl plane and is called the scattering vector. h , k , and l are the so-called Miller indices defining a unique plane of

reflection or diffraction spot. The planes are normally described in terms of their Miller Indices, which correspond to the intercept of the plane with the axes by $(a/h, b/k, c/l)$. Thus, for example, the (100) plane lies parallel to the bc face of the unit cell.

2.2.5 Reciprocal Lattice

Using Bragg's law, the orientation of the crystal in a diffracting position can be readily defined by the direction of the normal to the appropriate set of lattice points. Each diffraction spot is defined by two quantities, a magnitude and a direction. The magnitude and the direction combined forms a vector allowing each spot on a diffraction pattern to be intrinsically linked uniquely to a vector. The magnitude of the vector is proportional to d_{hkl} . Bragg's law shows that the diffraction angle is directly related to d_{hkl} : The equation below is true when $n=1$.

$$\sin \theta = (\lambda/2)(1/d_{hkl}) \quad (4)$$

The overall result of this is a set of vectors, one representing each reflection that is present in the diffraction pattern, whose end points define a lattice with dimensions inversely related to the dimensions of the crystal.

2.2.6 Structure Factor

The structure factor of a crystal is a mathematical description of how the crystal scatters incident radiation. The structure factor is a particularly useful tool in the interpretation of interference patterns obtained in X-ray, electron and neutron diffraction experiments. The intensity measured for a given reflection with Miller indices hkl is proportional to $|F(hkl)|^2$ where

$$F(hkl) = \sum_{j=1}^N f_j \exp[2\pi i(hx_j + ky_j + lz_j)] \quad (5)$$

f_j = the atomic scattering factor for X-rays for the j th atom with coordinates (x_j, y_j, z_j) expressed as fractions of the cell **a**, **b**, **c**. This is the structure factor equation in the absence of thermal motion or disorder.

The atomic scattering factor represents the scattering of each atom that is present in the unit cell. The strength of the scattering of individual atoms is related to the atomic number (Z). An increase in the Z value results in an increase in the scattering strength. The measured scattering intensity of X-rays in a diffraction experiment is proportional to Z^2 . With the increase of atomic number the difficulty of identifying the presence of lighter

atoms becomes increasingly difficult in the presence of heavier atoms as the scattering of the heavy atoms dominates the scattering of the lighter atoms.

The atomic scattering factor is modified in practice by several effects. X-ray photons interact with the electron cloud of an atom, however, these electron clouds are not points in space like the nucleus of an atom. These clouds possess a finite size of the same magnitude as the X-ray wavelength. The electrons are spread in space and consequently not all parts of the object are scattering in phase; the resulting destructive interference increases as the scattering angle increases and the scattering amplitude varies with 2θ . The atomic scattering factor (ratio of the amplitude scattered by an atom to that scattered by a single electron) thus falls off with $(\sin \theta)/\lambda$. A consequence of this is that the Bragg peaks at higher angles will generally exhibit a lower intensity compared to those seen at lower angles. A similar explanation is relevant to the effect of atomic thermal motion (see below)

A number of corrections have to be made to the structure factor for a standard experiment including corrections for thermal motion and scattering angle fall-off. For a structure with N atoms, each with atomic scattering amplitudes f_i and position \mathbf{r}_i in the unit cell, the structure factor for the Bragg reflections of index hkl is

$$F(hkl) = \sum_{j=1}^N f_j(\theta) T_j \exp[2\pi i(hx_j + ky_j + lz_j)] \quad (6)$$

T = temperature factor of the j th atom

The important information here lies in the scattering amplitudes $f_j(\theta)$ and in the temperature parameters T_j . As discussed above, the scattering factors $f_j(\theta)$ are strongly dependent on the scattering angle. At a scattering angle of 0° , f_j is proportional to Z , which is the atomic number of the scattering atom. As the scattering angle increases, the scattering factor $f_j(\theta)$ decreases. Thermal motion causes the scattering to fall off even more strongly, as it effectively increases the dimension of the scattering object, exacerbating the destructive interference effect. These factors combine to limit the resolution available in an experiment and make it difficult to determine the positions of light atoms accurately in the presence of heavy ones or to distinguish among heavy atoms that have similar atomic numbers. To help mitigate this by reducing the effect of thermal motion, single crystal diffraction experiments on molecular materials are frequently undertaken at low temperature unless there is a reason not to do so; most structures reported in this work have been collected at 100 K.

The application of a Fourier transform on an X-ray diffraction pattern makes it possible to determine the crystal structure of a system by transforming the diffraction pattern into an

electron density map. As the X-rays are scattered by the electrons it is actually the electron density of the crystal structure that is obtained from this process. The electron density for the unit cell can be expressed in terms of structure factors as:

$$\rho(xyz) = \frac{1}{V} \sum_{hkl} F(hkl) \exp[-2\pi i(hx + ky + lz)] \quad (7)$$

2.2.7. Structure Solution

2.2.7.1 The Phase Problem

Waves diffract from a primitive lattice of simple scatterers obeying Bragg's Law and this allows the determination of interplanar distances and thus the recovery of a description of the crystal lattice. However, the diffracted radiation suffers a phase shift arising from the spatial distribution of individual scatters. The amplitudes of the resulting structure factors are directly derivable from the experimentally measured intensities of the diffracted beams, but the phases are not. Without knowledge of the phases, it is not possible to reconstruct the individual atomic positions (with reference to the equation for the scattering density given above, this reflects the fact that $F(hkl)$ has both amplitude $|F(hkl)|$ which can be measured, and a phase, $\Phi(hkl)$, which cannot). Estimating the phases is an essential step in successful structure determination. This is known as the crystallographic phase problem.

2.2.7.2 Direct Methods for Structure Determination

This is the most widely used method for solving the phase problem in single crystal diffraction. It uses statistical relations between diffracted intensities to derive phase information. The pioneers of this were Karle and Hauptmann (Nobel laureates in 1985) [9]. The method depends on probabilistic relationships between phases for particular combinations of reflections. A typical triplet phase relationship in direct methods states that

$$\Phi(h_1, k_1, l_1) + \Phi(h_2, k_2, l_2) + \Phi(h_3, k_3, l_3) \approx 0 \text{ when } h_1, k_1, l_1 + h_2, k_2, l_2 + h_3, k_3, l_3 = 0 \quad (8)$$

Where $h_n, k_n, l_n = hkl$ reflection, and Φ are the phases.

This relation is more likely to be true for sets of stronger reflections. By establishing a large number of such triplets, and other related phase relationships, sets of approximate phases can be postulated, which can lead to successful structure solution. The success rate of these techniques is well over 90% for small/medium sized organic molecules.

2.2.7.3 Patterson Methods

This method produces peaks that correspond to vectors between pairs of atoms and the heavier the atoms are the larger the peaks will be. This helps to create a phase-less Patterson synthesis, based on the intensities of the reflections. The Patterson map contains peaks that are present at the distance vectors between the atoms present in the structure. The peak heights produced are proportional to the product of the number of electrons (Z) of the atoms located at the ends of the vector and the number of peaks are defined by $P=N^2-N$. P will become large very quickly and for the ease of interpretation Patterson methods are most effective when only a few heavier atoms are present; in chemical crystallography this relies on the molecular formula being appropriate, while for biological structures a heavy atom derivative can often be produced – only a few heavy atoms are required in the derivative. This method is very useful for solving structures that contain a small number of heavy atoms when the vectors relating these can be identified and analysed easily.

2.2.8 Difference Fourier Synthesis

The structural model obtained through structure solution may not be complete. Difference Fourier maps are used to show the electron density that has not been accounted for in the model. This is especially useful when searching for hydrogen atoms which appear as peaks in the Fourier difference maps of high resolution data. This can also highlight mis-assigned atoms or even atoms that are not modelled well. This process works by taking the structure factors calculated from the model away from those measured and using the differences with the phases calculated from the model, to produce an electron density map showing the unaccounted-for electron density.

$$\Delta\rho(xyz) = \sum_{hkl} (|F_{hkl}^o| - |F_{hkl}^c|) \exp(-2\pi i(hx + ky + lz) + \Phi_{hkl}^c) \quad (9)$$

There will be background noise in the difference map arising from Fourier termination errors, thermal disorder and electron density which is not taken into account in the atomic model, which includes bond density and lone pairs. These errors can be minimised by collecting data to as high a resolution as possible.

There will be some smearing present in the electron density map and the extent of this is determined by the thermal motion of atoms as they oscillate around their equilibrium positions. The structure is refined using simple ellipsoid models for these displacements, which in this way are allowed to be anisotropic, which gives a more accurate model of the

atomic motions in the crystal structure. This way of refinement does not take into account features including anharmonic thermal vibrations, non-spherical atoms and bonding density. The latter means that unaccounted density peaks normally arise between atoms especially in high resolution data sets when examining the difference Fourier maps.

2.2.9 Structure Refinement

Once a basic model, framework or fragment has been found using one of the solution methods available a combination of Fourier recycling and refinement is required to complete the model and finalise the determined structure.

For Fourier recycling, initially a Fourier map (electron density) is constructed using the calculated phases from the initial model from the structure solution, along with the measured reflection intensities. This Fourier map should reveal more structural information, including missing atoms, etc. The model of the structure is thus improved and this in turn improves the phases that are used for the next Fourier calculation, creating a cycle that if repeated will lead to a completed structure (usually meaning at this stage that all non-hydrogen atoms are included in the model). This Fourier recycling is an iterative process. Once the refinement process is in progress (see below), further Fourier map calculations are used to locate, if possible, the hydrogen atom positions and also as a check in the quality of the overall refinement.

Refinement is the method used for improving the description of the crystal structure once the approximate atomic positions are determined. In crystallography, this uses the method of least squares refinement, which involves varying the parameters in the model until the calculated structure factors from the refined model are in best agreement with the observed structure factor magnitudes. The parameters that are commonly varied are the positions (x , y , z) and thermal parameters (U_{ij}) of each of the atoms along with some overall parameters including the scale factor, absorption coefficient and some others. In the early stages of the refinement the non-hydrogen atoms are usually refined with an isotropic thermal factor (single parameter, assuming symmetrical vibration of the atom) but as the refinement proceeds these atoms can be refined anisotropically, with up to six thermal parameters U_{ij} , defining an ellipsoid reflecting the atomic motion,. Hydrogen atom thermal parameters usually have to be treated more cautiously, as isotropic, or sometimes tied to the values of the parent atom to which the hydrogen is bonded.

Least-squares refinement is the most common method used for the determination of inorganic and small-molecule structures. This minimises the function:

$$\sum w(Y_o - Y_c)^2 \quad (10)$$

where Y_o represents the observed structure factor, Y_c is the value calculated from the structural model, and w is an assigned weight. The weights will usually represent an estimate of the precision of the measured quantity. The sum is taken over all the measured reflections. For a pair of values of X and Y for measurements which are related by the equation $Y=mX+b$ a unique answer for the constants m and b can be obtained. In real problems such as a crystal structure determination, the error of the fit in the i th observation is $e_i=m_oX_i+b_o-Y_i$, where m_o and b_o are the best fit parameters. The principle of least squares states that the best-fit parameters are those that minimise the sum of the squares of the errors:

$$\sum_i e_i^2 = \sum_i (m_o X_i + b_o - Y_i)^2 \quad (i=2,\dots,N) \quad (11)$$

In crystal structure least squares refinement, there are always more observations than there are unknown quantities; the system is said to be overdetermined and this minimisation can be carried out. This is done in refinement programmes such as SHELXL [176]; in this work SHELXL was used as implemented in the WinGX program suite.

In addition to producing the best model parameters, the least squares refinement procedure will allow an estimate of the precision of the determined parameters. These are quoted as “standard uncertainties”. For example a fractional coordinate of 0.3452 might have a standard uncertainty of 0.0003, and this is expressed as 0.3452(3). Important derived geometrical parameters such as bond lengths can also be assigned a standard uncertainty reflecting their precision.

The number of parameters in a refinement is large, even for small molecules, and the overall quality of a refinement is evaluated by a series of residuals or “R-factors”. Two of these are most commonly used, R and wR2, defined as follows.

$$R = \frac{\sum ||F_o| - |F_c||}{\sum |F_o|} \quad (12)$$

Typical values of R of under 0.05 (5%) are normally anticipated in good quality small molecule structure determinations.

$$wR2 = \sqrt{\frac{\sum w(F_o^2 - F_c^2)^2}{\sum w(F_o^2)^2}} \quad (13)$$

The weighted R-factor incorporates weighting factors for the reflections based on the $\sigma(F^2)$ values.

A final check on the data refinement uses the goodness of fit (S), which shows how reliable the standard deviations of the positional and displacement parameters of the atoms are. This quantity is strongly influenced by the weighting scheme. The weighting scheme will be modified to force the S value to as close to 1.0 as possible. For a refinement on F^2 the S has the form:

$$S = [\sum [w(F_o^2 - F_c^2)^2] / (n-p)]^{1/2} \quad (14)$$

where n is the number of measured data and p the number of refined parameters.

2.3 Techniques and Instrumentation

X-ray diffraction techniques have become a big part of chemistry research and indeed the chemical industry, as they provide a way of determining a solid-state crystal structure with very good precision and high levels of reliability [114,115]. There are several different techniques that can be used, but the most commonly used techniques for the study of molecular structures are single crystal X-ray diffraction and powder X-ray diffraction.

2.3.1 Single Crystal X-ray Diffraction

One advantage to this type of technique is that it is non-destructive. If the crystals are not destroyed by air and moisture then they can be recovered. This means that extra studies can be carried out using the one crystal. X-ray diffraction can give vital information about the internal lattice of crystalline substances, including unit cell dimensions, and about the molecular structure, including bond lengths, bond angles, and details of site-ordering.

X-ray diffractometers (Figure 2.6) consist of three distinct parts: an X-ray tube, a sample holder and an X-ray detector. The X-rays are generated in a cathode ray tube by heating a filament to produce electrons, accelerating the electrons towards a target by applying a voltage, and the impact of the electrons with the target material produces characteristic X-ray spectra that are used for the diffraction experiment.



Figure 2.6. Two types of single crystal X-ray diffractometers used in this work. Left, a Bruker Apex II CCD diffractometer and right, a Rigaku R-axis/RAPID image plate diffractometer.

2.3.1.1 Selection of a Single Crystal

Selection of the best possible crystal is paramount to obtaining a good quality data set. The size of the crystal has an effect on the data collection. Small crystals in general result in only weak diffraction spots. The larger the crystal selected, the greater the scattering intensity will be, however, with increased volume there is also greater absorbance of X-rays by the crystalline material.

Once a good crystal has been found it is mounted onto a glass fibre fixed to a goniometer (Figure 2.7). This goniometer with the crystal is then placed onto the diffractometer and is centred in the X-ray beam. Since the crystal and the beam can both be very small it is important that the crystal is centred before any study takes place. The centring of the crystal will help to ensure that it remains in the centre of the beam throughout the data collection.



Figure 2.7. Goniometer on which a crystal is mounted and which allows centring of the crystal in the X-ray beam.

2.3.1.2 Data Collection

The first step in data collection is to determine the unit cell of the sample being investigated. This allows an initial estimation of the symmetry and the space group of the crystal to be made. The advantage of carrying this out is that the process is very quick and a crystal suitable for full data set collection can be found whilst others that are, for example, known starting materials or previously determined products, can be eliminated. These initial runs consist of short scans with the length of exposure being adjusted so that the exposure is increased for weakly diffracting crystals. This part of the process is crucial as the selection of a crystal system with too high symmetry, if wrong, can result in a data collection strategy that does not collect all the necessary reflections. If there are any doubts about the crystal system, selecting a triclinic data collection will ensure a complete dataset is obtained. A data collection strategy is calculated to find the most efficient way to collect a complete dataset. A full data collection can take anything from a few hours to several days to collect depending on the crystal system and the strength of scattering from the crystal.

Upon completion of the data collection, hopefully, all the necessary reflections have been collected. Each of the reflections collected have a set of Miller indices, an intensity, and an estimation of how precise the measurement of intensity is. Corrections are applied to the data during data reduction such as an absorption correction.

The reflection positions and intensities are processed using the software related to the specific diffractometer used. During the course of the research three diffractometers were used and these are summarised in Table 1.1, The processed data are then transferred to programs such as WinGX [116] or CRYSTALS [117] for structure solution, refinement and visualisation. WinGX has been used for the structure determinations reported in this thesis.

Table 2.1 Diffractometers used and associated software used

Diffractometers Used	Software used to process the data
Bruker APEX-II CCD	Bruker APEX2 (2009) [118]
Rigaku R-Axis/RAPID Image Plate	CrystalClear 1.4.0 (Rigaku, 2008) [119]
Bruker Nonius Kappa CCD	Collect (Bruker AXS BV, 1997-2004) [119] HKL Scalepack (Otwinowski & Minor 1997) [120] HKL Denzo and Scalepack (Otwinowski & Minor 1997) [120]

2.3.2 Powder X-ray Diffraction

Powder X-ray diffraction is valuable both in combination with single crystal diffraction and also as a stand alone technique. Powder diffractometers can be used to gain information about the composition of bulk materials for example if there are multiple phases, and to determine if a crystal that has been studied via single crystal diffraction is representative of the bulk of the sample. Information about unit cell size and symmetry and in some cases full 3-dimensional structural information comparable to that obtained from single crystal methods can be obtained. Powder diffraction can also be used conveniently on samples under the influence of external factors such as temperature, pressure or applied magnetic field, under laser illumination of a sample, and even as a function of time or chemical environment during the synthesis of materials, giving very valuable kinetic and mechanistic insight into their formation.

Like single crystal diffraction, the peak positions are governed by the size, shape and symmetry of the unit cell. The intensities of the peaks are determined by the arrangement of scattering density (atomic coordinates) within the unit cell. The shape of the peaks includes the convolution of instrument parameters (source, optics and detector contributions) and information about the microstructure (domain size, strain) of the sample.

2.3.2.1 Experimental Procedure

There are two common types of powder diffractometer geometries: flat-plate where the X-ray beam is reflected from the sample; and capillary where the X-ray beam is transmitted through the sample. In both methods, a detector is used to scan through the scattering angle (2θ). A powder diffraction sample contains a very large number of small single crystals (crystallites) in random orientations, and the diffraction pattern obtained is therefore a compression of a three-dimensional diffraction pattern into a spatially averaged

one-dimensional diffraction pattern. Preferred orientation can occur when a sample is in a capillary or a flat plate and the orientation of each individual microcrystal is preferential in some directions. The effect on the diffraction pattern is to superficially enhance the intensities of the Bragg peaks of some reflections and decrease others. This arises due to a greater tendency for the crystallites in the powder sample to be orientated more frequently in one specific way, or one set of ways, rather than all of the others. For example if the crystal form is that of needles, they will tend to align in one direction along a capillary. This can arise if the powder sample has not been sufficiently ground. A method of reducing this effect is to spin the capillary or by use of the flat-plate. On the other hand, for crystals with a very plate-like morphology, often flat-plate geometry can result in problems with preferred orientation, so care has to be taken.

2.3.2.2 Phase Identification

There can be several phases present in a bulk sample being analysed. Each phase gives rise to a characteristic set of peaks in a powder diffraction pattern (a fingerprint). These patterns can be compared to a database of known diffraction patterns or compared against patterns calculated from single-crystal diffraction studies. Many single crystal refinement packages or related programs contain an option to calculate a powder pattern from either a refined structural model or directly from patterns calculated from the atomic coordinates available from single-crystal diffraction studies. There are many examples where the few single crystals produced in a synthesis are due to minor products from side reactions or impurities. The presence of crystalline co-products can be readily identified. Powder diffraction is a fast technique that can shed considerable light on otherwise conflicting pieces of analytical data.

2.3.2.3 Quantitative Analysis

Powder diffraction can be used to obtain quantitative information about the composition of a multiphase sample from powder diffraction data. Many techniques have been developed based on the analysis of intensities of individual peaks due to different phases contributing to the pattern, on whole pattern intensity analysis, or on multiphase Rietveld refinement. One such program is PolySNAP which uses cluster analysis to quantitatively analyse the compositions of multiple samples [121].

Extreme care must be taken when determining and interpreting quantitative composition. The end results can be severely influenced by methods of sample preparation, data

collection and analysis. The calibration of the equipment used must be carefully carried out before the experiment. It is possible to estimate the quantity of amorphous material in a sample by powder diffraction by the careful quantitative dilution of a powdered sample with an additional crystalline phase.

2.3.2.4 Intensity Information

The intensities of the peaks in a powder diffraction pattern contain information about atomic coordinates and displacement parameters, just as in a single-crystal experiment. In the early days, structural work using powder diffraction data analysed extracted intensities and refinement methods essentially identical to those used in single crystal work. Nowadays it is more common to employ whole-pattern fitting methods to extract structural information – principally the Rietveld method [122].

2.4 Crystallisation

There are many different types of methods that can be used for crystallisation attempts. The main ones are slow evaporation [123], slow cooling, vapour diffusion [124] and solvent diffusion [125]. The method used depends on the properties of the components, and the size of the single crystals that may be required (Figure 2.8).

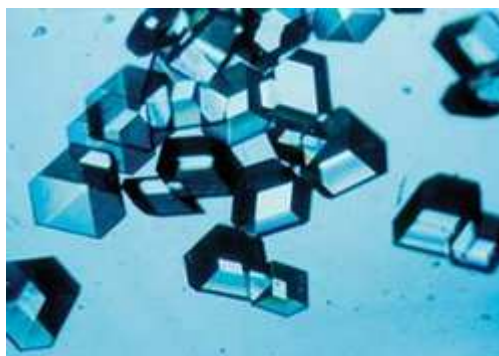


Figure 2.8. Typical single crystals used for diffraction; these are of insulin.

2.5.1 Crystallisation by Slow Evaporation

A small quantity (few mg) of the sample is placed into a sample vial and solvent is added to fully dissolve the contents of the vial. A plastic clip on lid is placed over the top of the vial and a few holes made in this to allow the sample to evaporate slowly and hopefully, allow crystals to form (Figure 2.9). The temperature can be controlled at both elevated (using a hot plate) and reduced (using a cold room) temperatures. The hot plate temperature can be varied depending on the solvent being used.

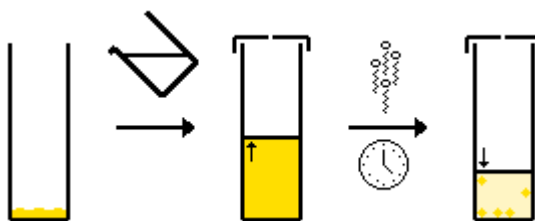


Figure 2.9. Schematic of slow evaporation crystallisation.

2.5.2 Solvent Diffusion

Solvent diffusion is another technique that can be used to grow crystals. This technique involves dissolving the sample in a solvent in which it is readily soluble (good solvent). Another solvent is then added to this sample vial in which the sample will not be readily soluble (poor solvent). The two solvents used must be co-miscible. The poor solvent is added using a syringe and is added slowly. This helps to make sure that there are two layers present. If the poor solvent is denser than the good solvent it is added to the bottom of the sample vial and if the poor solvent is less dense it is added to the top of the sample vial. These two layers then mix slowly and the crystals hopefully grow at the interface. If necessary, the vial can be cooled to slow the diffusion rate and to reduce solubility. Figure 2.10 shows the procedure with the yellow solvent being the good solvent, the blue solvent being the poor solvent and the green area being the interface.

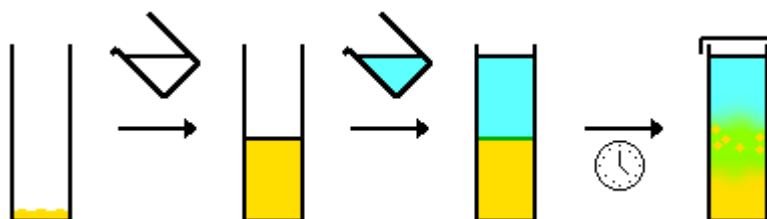


Figure 2.10. Schematic of solvent diffusion crystallisation.

2.5.3 Vapour Diffusion

The technique can also be known as isothermal distillation. The setup is as shown in Figure 2.11.

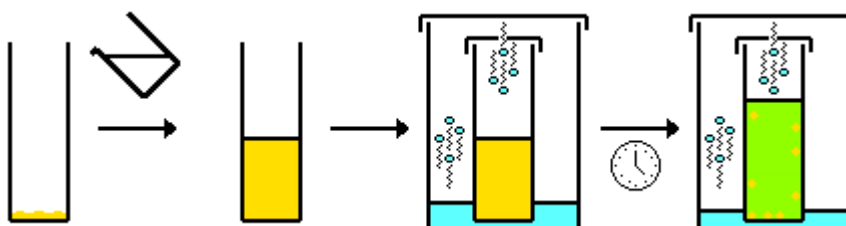


Figure 2.11. Schematic of vapour diffusion crystallisation.

The yellow solvent is the good solvent and the blue solvent is the poor solvent. The good solvent in which the material is dissolved is placed into a sample vial. This sample vial is then placed into a bigger vial containing the desired poor solvent. The larger vial is then covered. The poor solvent diffuses through the vapour phase into a solution of the target compound in the good solvent and thus reduces the solubility. The main reason to use this technique is that it has slow rate of diffusion, has high controllability and is very adaptable.

2.5.4 Variable Temperature Parallel Crystallisation –The Microvate

This piece of equipment can be used to carefully control the temperature of the crystallisations, and allow many crystallisations to be set up simultaneously. The Microvate has 48 wells, arranged in rows of four (Figure 2.12), with each of the 12 rows of four able to be independently temperature-controlled. The Microvate can be used to heat samples up to a maximum of 165°C at a controlled rate. Samples are normally held at a fixed temperature for a period of time and then cooled down at a slow rate to try and induce crystals to form.



Figure 2.12. The Microvate parallel crystalliser.

2.6 Differential Scanning Calorimetry

Differential Scanning Calorimetry or DSC has primarily been used in this thesis to measure the temperature of melting of the sample that has been produced through a co-crystallisation attempt. In this context, it has been used as an indicator of successful co-crystallisation where a different melting point for the sample from those of the starting materials can indicate the formation of a molecular complex. It has also been used as an indicator of purity, where a pure product would only be expected to show a single melting point. Other phase transitions can also be identified during a DSC experiment such as glass transitions and crystallisation events (Figure 2.13).

A sealed pan with the sample contained within it and an empty reference pan are placed into the sealed chamber used for heating the sample. Both the pans are enclosed in a sealed furnace and this is heated over a temperature range specific to the sample of interest. The rate of heating is kept constant. The difference in the energy required to maintain the two pans at close to identical temperatures is measured by the heat changes in the target sample. Whilst the sample is being heated, if there is an exothermic peak in the resultant thermogram, this can be classed as a crystallisation peak and the peak appears as a maximum in the graph. When the sample reaches its melting temperature an endothermic peak appearing as a minimum in thermogram is observed. Both of these phenomena can be seen in a single experiment as the sample is heated it will reach its melting point melt resulting in a melting peak and if the user has programmed the sample to be cooled back to room temperature a crystallisation peak may be evident as the sample cools.

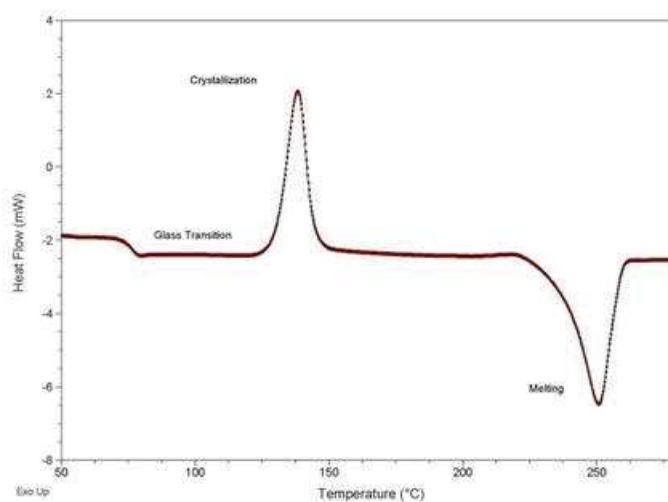


Figure 2.13. A typical DSC thermogram showing the different types of thermal events.

2.7 Thermogravimetric Analysis (TGA)

In the TGA technique, the sample is heated in a closed furnace in a similar fashion to the DSC, however, the sample is not compared to a reference sample but instead the change in weight of the sample is measured as the temperature of the sample is slowly increased. This technique therefore monitors the mass of a substance as a function of time. One main use for this technique is to investigate solvent loss from a sample and in particular whether this occurs before the sample melts. Such solvent loss results in a step in the thermogram prior to decomposition of the sample (Figure 2.14).

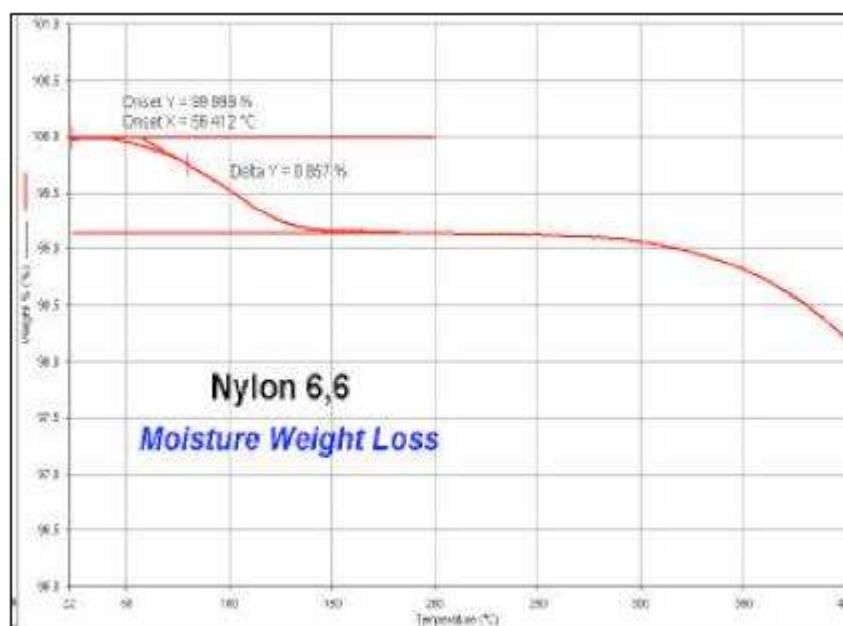


Figure 2.14. A TGA thermogram showing a step as water is released from the sample before the decomposition temperature is reached[126].

2.8 dSNAP

The dSNAP program was a tool that was first released for use in 2005 [127]. This program aims to make it possible to easily analyse the large quantity of data contained within the CSD. Searching for even a simple fragment or molecule can result in a large number of hits in a CSD search and the analysis of these results can be time consuming and difficult. The dSNAP program was introduced to help the structural chemist extract information from CSD searches making it quick and easy to compare these results with not only other hits but also the users own work to see where their results fit in with what has been previously published.

The procedure of dSNAP is as follows and will show the complicated mathematics to the program. The user will define a search in the CSD and search accordingly. The three dimensional option has been utilised in the CSD when conducting the search meaning all the interatomic distance and angles have been defined. The number of geometrical parameters is equal to

$$(l/2)(L-1)+(l/2)(L-1)(L-2)=l/2(L-1)^2 \quad (15)$$

Where l is equal to the number of atoms. There will be obvious redundancy here and the user can also eliminate hits from the list before submission to the dSNAP program. Once this process is complete the data matrix will be inserted into the dSNAP program along with the corresponding files that have been output from the search conducted. The files are essential to the program running and must be present. The data matrix will be converted to a symmetric (n x n) Minkowski distance matrix, d^s , via the standard formula:

$$d_{ij}^s = (\sum_{k=1}^m w_k |x_{ik} - x_{jk}|^\lambda)^{1/\lambda} \quad (16)$$

The λ is a user selectable parameter in dSNAP that will have a default value of 2 corresponding to a Euclidean distance matrix. A value of 1 can also be used occasionally and this refers to the traditional city-block distance. The superscript 's' for the d matrix implies that in the subject or stimulus space to distinguish it from the related space. The w present refers to the weights. The weights are user selectable parameters with the default of unity. One of the most useful parts of the dSNAP program is the metric multidimensional scaling of MMDS. This is used to generate a three-dimensional Euclidean space in which each point will represent one fragment of the data set making comparisons between separate fragments easy to make. So from the d^s a matrix A(n x n) can be constructed:

$$A = -\frac{1}{2}(I_n - \frac{1}{n}i_n i_n') D^s (I_n - \frac{1}{n}i_n i_n') \quad (17)$$

From the equation the I_n is an (n x n) identity matrix, i_n is an (n x 1) vector unities and D^s is a matrix of squared distances. This means that the eigenvectors of A, v_1, v_2, \dots, v_3 form a

vector V and the corresponding $\lambda_1, \lambda_2, \dots, \lambda_n$ give a second vector Λ . A total of p eigenvalues are positive and the remaining $(n - p)$ are set to zero. A set of coordinates in p dimensions can be defined via the matrix $X(n \times p)$ when

$$X = V \Lambda^{1/2} \quad (18)$$

This means that p can now be set to 3 and the work can be done in three dimensions. The X matrix can be used to plot each data set fragment as a single point in the three dimensional plot.

For the user the program takes an output from a CSD search. The fragment under investigation can be a whole molecule under which some geometrical flexibility is anticipated or a fragment that contains hydrogen bonds between two molecules in a crystal structure. The search results are input into the dSNAP program which analyses both the bonded and non-bonded distances and the angles within the fragments using statistical methods and subsequently groups the results into clusters with each cluster comprising hits that are closely related.



Figure 2.15. Difluoroalkene used in analysis using the dSNAP program [127].

This can be illustrated using the example that was published around the time of the first release of the program, involving a difluoroalkene [127] (Figure 2.15). This difluoroalkene is set up as the fragment for analysis and a CSD search for this fragment resulted in 58 hits from 33 crystal structures. This means that some of the crystal structures will have more than one hit per structure. The hits are analysed to compare the distances and angles of each individual hit and to separate these hits out into clusters made up of hits that are similar in nature. After the initial analysis, the main dSNAP window will appear as shown in Figure 2.16 (left). This part will show the individual hits that have been analysed with the CSD ref code appearing below a coloured circle. All of the hits involved in the

calculations will be represented by a unique circle. The colour of this circle shows, which cluster it is a member of and this colour will be maintained in all of the outputs possible.

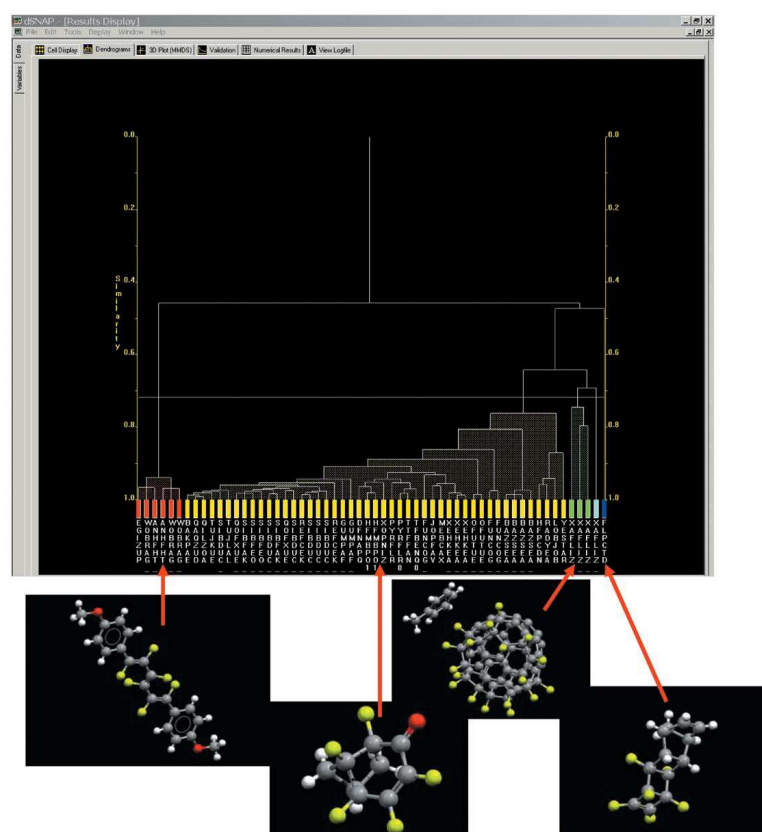


Figure 2.16. Dendrogram for the analysis of acid-amide interactions [127].

The most useful initial visualisation of the dSNAP output is the dendrogram (Figure 2.16, right); this is produced from dSNAP and shows tie bars that link the individual hits together, with the lower the tie bar linking them, the more similar are the hits. The dendrogram can be difficult to understand when there are a large number of hits present and using the MMDS plot can help to make it easy to visualise how similar some of the hits are. The MMDS plot colours the hits by cluster (Figure 2.17). In this example, the clusters are well-defined with the clusters forming closely packed units and each cluster is well separated from the others. However, if the clusters are not tightly packed or if more than one cluster merges together the cut-level for the dendrogram may be incorrect and this can be adjusted by the user to the level deemed correct. dSNAP does not provide any chemical interpretation for the clustering.

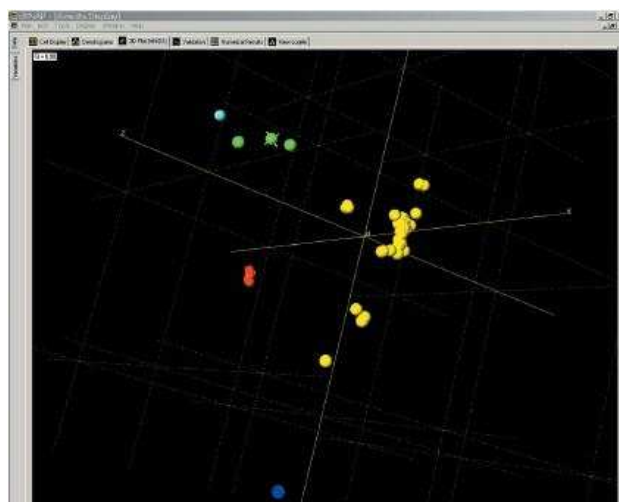


Figure 2.17 MMDS plot showing the separate clusters and how closely related they are [127].

In this example the hits were clustered into 5 separate clusters (A-E) with the calculated cut level of 75%. The clusters need to be analysed by the user to determine if the clustering has worked. In this case the clusters were separated into: A (red) contain all the *trans*-fluorine fragments, and groups B-E contain the *cis*-fluorine fragments. Group B (yellow) contains what can be commonly considered, the normal *cis*-fluorine structures. The clusters C (green) and D (light blue) contain a total of four fragments that all come from the same structure, which contains an aromatic carbon-cluster with F bound to the surface. The four fragments each contain a carbon atom that has four groups bond to it, with one being a double bond, and as a result are not chemically reasonable. The final cluster contains only one hit and the fragment is part of a cyclobutane ring and as a result of this has a more strained chemical environment.

The MMDS plot helps to show the effectiveness of the clustering as the hits are well spaced out and shows the clustering has been correctly done. The green and light blue clusters are very close in this case but this would be expected due to the obvious similarities.

Chapter 3. Molecular Complexes of Cytosine with Benzoic Acid and Its Fluorinated Derivatives

3.1. 1:1 Molecular Complex of Cytosine and Benzoic acid [128]

Cytosine and benzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.3mm x 0.4mm x 0.2mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. The crystal structure determined at room temperature has been reported previously [128] but data were re-collected and the structure solved at 100K to allow direct comparison with the other molecular complexes reported in this thesis. Crystallographic data are summarised in Table 3.11, at the end of this Section 3.5. The molecular complex forms in a 1:1 ratio of cytosine to benzoic acid. There are four molecules in the asymmetric unit, and the crystal belongs to the space group $P\bar{1}$ with eight molecules in the unit cell. Whilst the ratio in the product is expressed as 1:1, one cytosine has gained a hydrogen from the carboxyl group of a benzoic acid molecule, essentially creating four different molecular species, two of which carry a charge: a neutral benzoic acid and an deprotonated benzoic acid (benzoate); and a protonated cytosine (cytosinium) and one neutral cytosine (Figure 3.1).

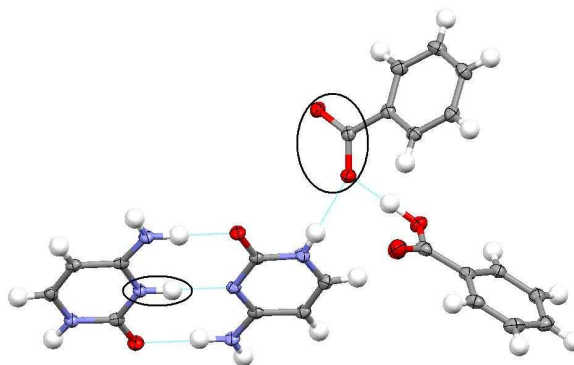


Figure 3.1. The asymmetric unit of the cytosine benzoic acid molecular complex showing the four independent molecules.

The proton transfer allows the formation of a pseudo-Watson-Crick base paired dimer between the neutral and protonated cytosine molecules. This pseudo-Watson-Crick base pair has three hydrogen bonds forming the supramolecular unit, one of which is charge assisted (Figure 3.2).

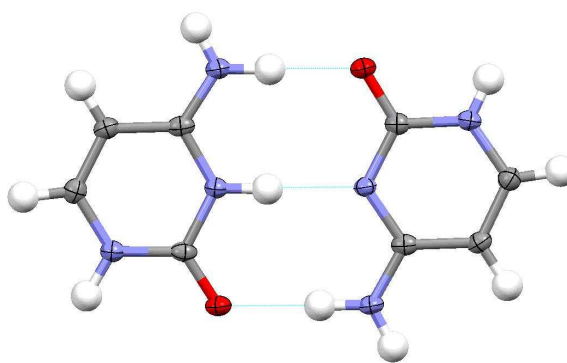


Figure 3.2. The pseudo-Watson-Crick base cytosine-cytosinium pair found in the 1:1 molecular complex of benzoic acid and cytosine.

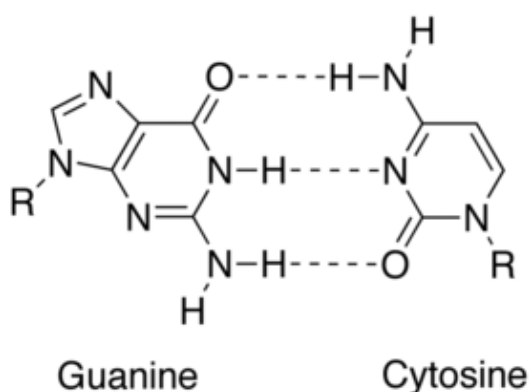


Figure 3.3. Conventional Watson-Crick base pair involving cytosine and guanine.

Figure 3.3 illustrates conventional Watson-Crick base pairing between cytosine and guanine. In general, these pairs are constructed of three hydrogen bonds and this is similar to the cytosine-cytosinium unit found in the cytosine benzoic acid molecular complex. This configuration is possible between guanine and cytosine due to the positioning of the hydrogen atoms on both molecules, guanine having two donor groups and one acceptor and cytosine having two acceptor groups and one donor. Cytosine would not be able to form the pseudo-Watson-Crick with itself in the absence of hydrogen transfer. The only way that cytosine can form this type of base pair is through hydrogen transfer to one molecule and not the other; this allows the three hydrogen bonds to form instead of the normal Hoogsteen unit that is observed in the crystal structure of cytosine itself [129].

The hydrogen bonds in the cytosine base pair are all of moderate strength [130]. However, all three hydrogen bonds in the pseudo-Watson-Crick unit are shorter than those found in the molecular complex of ethyl-guanine and cytosine [131]. This is a consequence of the central hydrogen bond being charge assisted.

Table 3.1. The hydrogen bonds found within a conventional Watson-Crick unit of the 9-ethylguanine 1-methylcytosine molecular complex (EGMCYT10 [131]).

Hydrogen Bond	D...A Distance
N3-H---O1A	2.934(2) Å
N1---H-N1A	2.924(2) Å
O1---H-N3A	2.806(2) Å

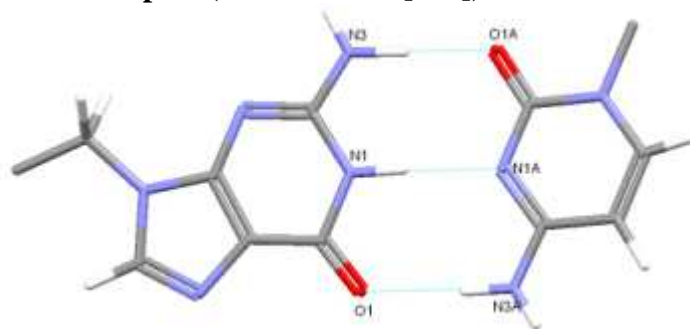


Table 3.2. The hydrogen bonds found within the pseudo-Watson-Crick cytosine-cytosinium unit of the molecular complex of cytosine with benzoic acid.

Hydrogen Bond	D...A Distance	D-H Distance	D-H...A Angle
N3-H---O1A	2.914(2) Å	0.98(2) Å	174.2(16)
N1---H-N1A	2.844(2) Å	0.98(2) Å	177.5(18)
O1---H-N3A	2.771(2) Å	0.99(2) Å	178.2(16)

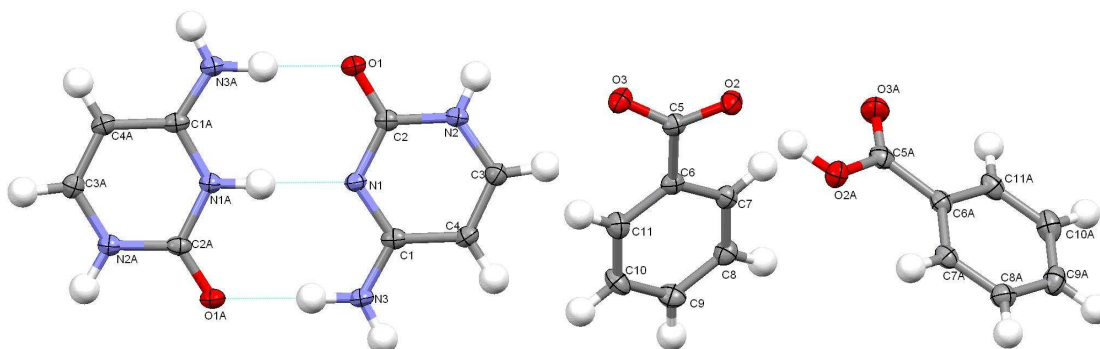


Figure 3.4. The labelling scheme for the cytosine molecules (left) and the benzoic acid molecules (right) in the cytosine and benzoic acid molecular complex.

3.1.1 Benzoic Acid Bond Distances

Table 3.3 illustrates that the proton transfer has affected the bond distances of both the carboxylate group of the benzoate molecule and the ring of the cytosinium molecule. When the benzoic acid and the benzoate molecules are compared, it becomes obvious that the proton transfer has had an affect. In the benzoic acid, the carboxyl group behaves like a neutral benzoic acid group would normally. The carboxyl group is made up of a typical C=O in the C5A-O3A bond (1.215(3) Å), and a typical C-OH which exhibits a standard single C-O bond distance in the bond C5A-O2A (1.330(3) Å).

Once the proton is lost from the carboxyl group a carboxylate group is produced and the bond distances change to help balance the charge now present. The C-OH bond changes to a C-O⁻ bond, represented in the structure by the C5-O2 bond distance. This distance has

shortened by approximately 0.05 Å with a bond length of 1.275(3) Å. The C=O, represented in the structure by C5-O3, has also altered, lengthening from a typical C=O bond distance by approximately 0.040 Å. This means that both bond distances are intermediate between typical single or double bonds and this points to the fact that the charge is delocalised over the carboxylate group.

3.1.2 Cytosine Bond Distances

The cytosine molecule undergoes a bigger change than that of the benzoic acid molecule. The formation of a cytosinium molecule causes five bonding distances to change notably, as shown in Table 3.3. These changes reflect the fact that the charge left on the cytosine molecule is spread over the side of the cytosine molecule that is involved in the base pair. This leads to both bond shortening and lengthening, and some consequent changes in the ring angles.

Table 3.3 Bond distances in each molecule present in the molecular complex of cytosine and benzoic acid

Cytosinium				Cytosine			
Bond length (Å)		Angle (°)		Bond length (Å)		Angle (°)	
O1A-C2A	1.238(3)	C1-N2-C2	124.0(2)	O1-C2	1.251(3)	C1A-N2A-C2A	119.3(2)
C2A-N2A	1.359(3)	N2-C2-N3	116.3(2)	C2-N2	1.371(3)	N2A-C2A-N3A	119.3(2)
N2A-C3A	1.362(3)	C2-N3-C3	121.8(2)	N2-C3	1.360(3)	C2A-N3A-C3A	121.8(2)
C3A-C4A	1.343(3)	N3-C3-C4	122.2(2)	C3-C4	1.342(3)	N3A-C3A-C4A	120.4(2)
C4A-C1A	1.420(3)	C3-C4-C1	118.0(2)	C4-C1	1.424(3)	C3A-C4A-C1A	117.3(2)
C1A-N3A	1.312(3)	C4-C1-N2	117.8(2)	C1-N3	1.328(3)	C4A-C1A-N2A	121.8(2)
C1A-N1A	1.361(3)			C1-N1	1.348(3)		
N1A-C2A	1.374(3)			N1-C2	1.356(3)		
Benzoic acid				Benzoate			
Bond Length (Å)		Angle (°)		Bond Length (Å)		Angle (°)	
O2A-C5A	1.330(3)	C6A-C7A-C8A	120.5(2)	O2-C5	1.275(3)	C6-C7-C8	120.0(2)
C5A-O3A	1.215(3)	C7A-C8A-C9A	119.7(2)	C5-O3	1.253(3)	C7-C8-C9	120.2(2)
C5A-C6A	1.489(3)	C8A-C9A-C10A	119.7(2)	C5-C6	1.506(3)	C8-C9-C10	120.1(2)
C6A-C7A	1.392(3)	C9A-C10A-C11A	120.6(2)	C6-C7	1.389(3)	C9-C10-C11	119.8(2)
C7A-C8A	1.384(3)	C10A-C11A-C6A	119.7(2)	C7-C8	1.385(3)	C10-C11-C6	120.3(2)
C8A-C9A	1.392(3)	C11A-C6A-C7A	119.5(2)	C8-C9	1.383(3)	C11-C6-C7	119.5(2)
C9A-C10A	1.387(3)			C9-C10	1.390(3)		
C10A-C11A	1.381(3)			C10-C11	1.381(3)		
C11A-C6A	1.395(3)			C11-C6	1.396(3)		

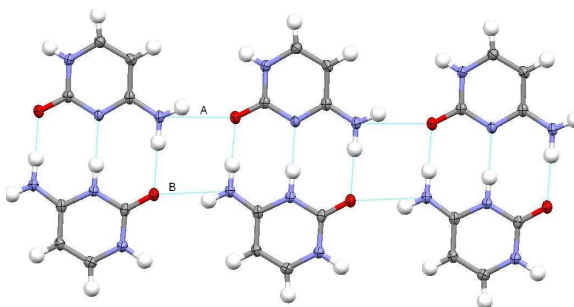


Figure 3.5. Cytosine homo-base pairs forming into a chain along the *b*-axis in the molecular complex of cytosine and benzoic acid.

The pseudo-Watson-Crick cytosine dimers form chains along the *b*-axis with two N-H...O hydrogen bonds connecting one base paired unit to another (Figure 3.5). These hydrogen bonds are of moderate strength with A being of length 2.860(2) Å and B of length 2.876(2) Å and form between the carbonyl groups and the amine groups of like molecules (i.e. cytosine to cytosine or cytosinium to cytosinium). Hydrogen bonded rings are thus formed of type $R^2_4(8)$. The chains are planar.

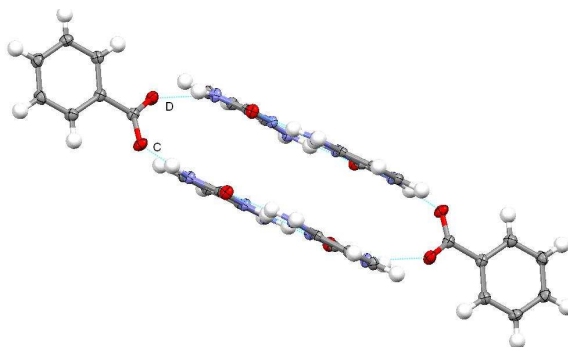


Figure 3.6. Benzoate molecules connecting the layered cytosine chains along the *b*-axis in the molecular complex of cytosine and benzoic acid.

The benzoate molecules link parallel cytosine chains to one another through two N-H...O hydrogen bonds of moderate strength and length C 2.666(2) Å and D 2.821(2) Å (Figure 3.6). This six molecule building block (Figure 3.6) stacks along the *c*-axis. In pure benzoic acid, the carboxylic acid group lies in the plane of the benzene ring [128]. This is also the case for the neutral benzoic acid molecule in this complex. The benzoate molecule, however, is not planar and the carboxylate group lies at an angle of approximately 25° with respect to the benzene ring. This twisting is a consequence of the connecting role of this molecule such that it can form two hydrogen bonds to two different layers of cytosine chains.



Figure 3.7. Twisting of the carboxylate group in the benzoate molecule in the molecular complex of cytosine and benzoic acid.

The neutral benzoic acid interacts with the deprotonated benzoic acid through a single O-H...O hydrogen bond with an O...O distance of 2.567(2) Å to the slightly longer C-O group of the carboxylate group. It is therefore charge assisted in nature and is the shortest hydrogen bond observed in this structure. This is the only hydrogen bond in which the neutral benzoic acid participates.

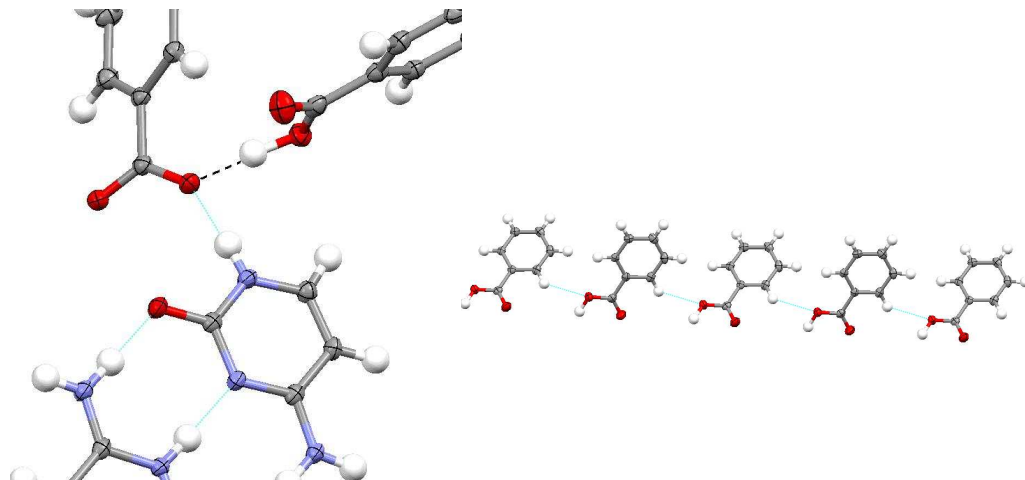


Figure 3.8. Neutral benzoic acid hydrogen bonded to the deprotonated benzoic acid in the molecular complex of cytosine and benzoic acid (left) and the benzoic acid chain linked via weak O...H-C hydrogen bonds (right).

The twisting of the carboxylate group allows for some $\pi\cdots\pi$ interactions to form connecting a neutral benzoic acid with a charged benzoate. This leads to layers of π stacked benzoic acid molecules. The benzoic acid chains, which form along the *c*-axis, are predominately held together by weak interactions. The neutral benzoic acid forms weak O...H-C hydrogen bonds to create a chain of benzoic acid molecules, the chain is made up of only neutral benzoic acid molecules. This connects the six molecule units to each other via weak interactions and these interactions occur along the *c*-axis.

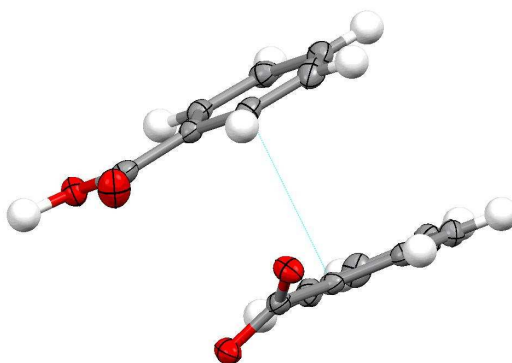


Figure 3.9. $\pi\cdots\pi$ stacking interactions in the molecular complex of cytosine and benzoic acid

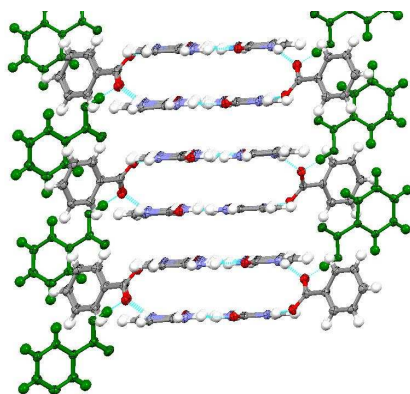


Figure 3.10. Link between separate six molecule units via weak interactions in the molecular complex of cytosine and benzoic acid. The molecule involved in weak interactions between both blocks is represented by green.

Figure 3.11 shows that there is prominent base-stacking in the structure. There are no direct hydrogen bonds formed between chains of parallel cytosine molecules, but these chains stack on top of one another. There are two cytosine stacking interactions in this molecular complex. The first is between the two layers within a six molecule unit (Figure 3.6). The cytosinium molecules (red in Figure 3.11) stack on top of the neutral cytosine molecules (green in Figure 3.11). The approximate spacing between these two layers is 3.01 Å. The stacking interactions between different six molecule units also show neutral cytosine molecules stacked upon charged cytosinium molecules but the layer spacing is slightly larger being approximately 3.2 Å.

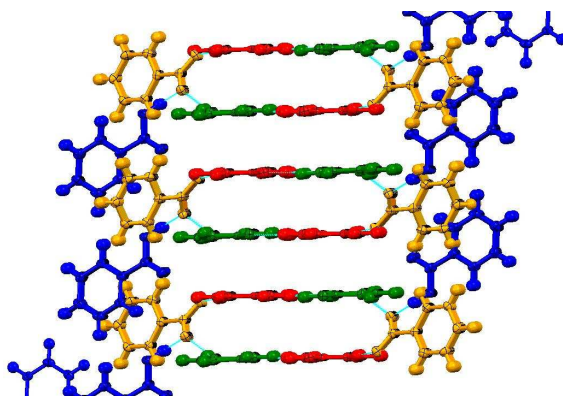


Figure 3.11. Stacking interactions in the molecular complex of cytosine and benzoic acid between building blocks with colour scheme: red (charged cytosine), green (neutral cytosine), orange (charged benzoic acid) and blue (neutral benzoic acid)

3.2. 1:1:0.5 Molecular Complex of Cytosine and 2-Fluorobenzoic acid Hemihydrate

Cytosine and 2-fluorobenzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.2mm x 0.2mm x 0.2mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 3.11.

The molecular complex forms in a 1:1 ratio of cytosine to benzoic acid in a similar manner to that of the benzoic acid molecular complex but with water also incorporated into the structure (1:1:0.5, cytosine: 2-fluorobenzoic acid: water). A hydrogen atom from the carboxyl group of the 2-fluorobenzoic acid is again transferred to a cytosine creating five independent molecules in the asymmetric unit (Figure 3.13). Both the 2-fluorobenzoic acid and 2-fluorobenzoate molecules show orientational disorder with respect to the fluorine atom. This disorder was modelled with respect to two separate components involving a fluorine atom, two carbon atoms and a hydrogen atom (Figure 3.12). The major component was refined to have an occupancy of 97% and the minor component of 3%.

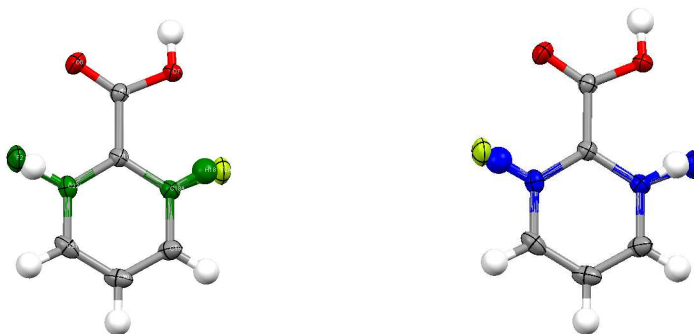


Figure 3.12. The major (green) and minor (blue) components of the 2-fluorobenzoic acid molecules in the molecular complex of cytosine and 2-fluorobenzoic acid hemihydrate.

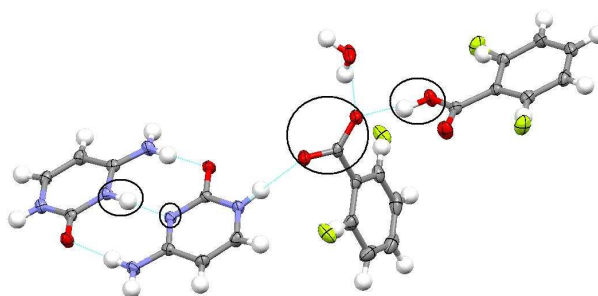


Figure 3.13. The asymmetric unit showing proton transfer from one benzoic acid molecule to one cytosine molecule in the molecular complex of cytosine and 2-fluorobenzoic acid hemihydrate.

The proton transfer again allows the formation of a pseudo-Watson-Crick base paired dimer between the neutral and protonated cytosine molecules.

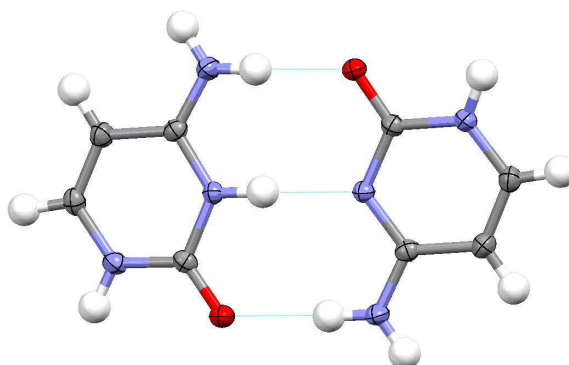


Figure 3.14. The pseudo-Watson-Crick base pair observed in the 1:1 molecular complex of cytosine and 2-fluorobenzoic acid hemihydrate.

Table 3.4. Hydrogen bond distances of the pseudo-Watson-Crick hydrogen bonding motif in the molecular complex of cytosine and 2-fluorobenzoic acid hemihydrate.

Hydrogen Bond	D---A Distance
N1-H...O1A	2.836(2) Å
N2...H-N2A	2.860(2) Å
O1...H-N1A	2.912(2) Å

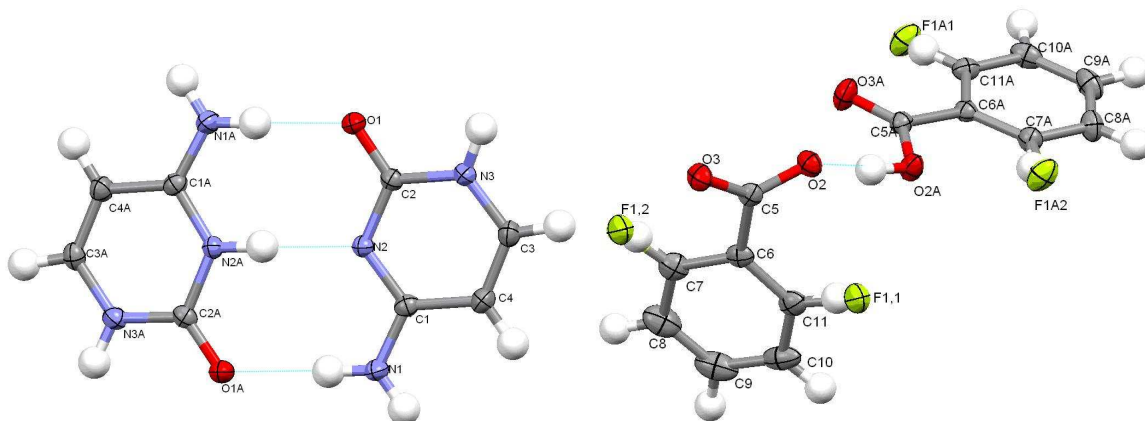


Figure 3.15. The labelling scheme for the cytosine molecules (left) and the benzoic acid molecules (right) in the molecular complex of cytosine and 2-fluorobenzoic acid hemihydrate

This structure also produces a two hydrogen bonded link between base-pairs. In the benzoic acid cytosine molecular complex, the link was planar and more square in shape; here it is stepped and in the shape of a diamond with a large asymmetry in the hydrogen bond lengths. The same two hydrogen bonds make up the link with one O...H-N bond and one N-H...O bond. The first of these hydrogen bonds is from the amine group of the neutral cytosine to a nearby oxygen on another neutral cytosine molecule. This hydrogen bond is of moderate strength and has an N...O distance of 3.021(2) Å. The other hydrogen bond making up the second link is from the amine group of a cytosinium molecule to the oxygen

of a nearby cytosinium molecule. This hydrogen bond is of weak strength and has an N...O distance of 3.233(2) Å; the primary hydrogen bond involving this amine group is to a 2-fluorobenzoic acid molecule and is of N...O distance 2.897(2) Å (see below). This link is shown in Figure 3.16. Both these hydrogen bonds deviate significantly from linearity having N-H...O angles of 140(3)° and 118(3)°, respectively. The pseudo-Watson-Crick base pairs are therefore stepped relative to one another within the chains (Figure 3.17).

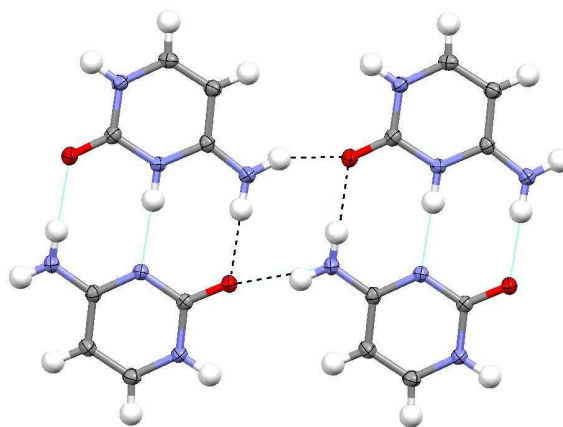


Figure 3.16. Base paired chain linked by hydrogen bonds making a diamond shape in the cytosine 2-fluorobenzoic acid hemihydrate molecular complex.

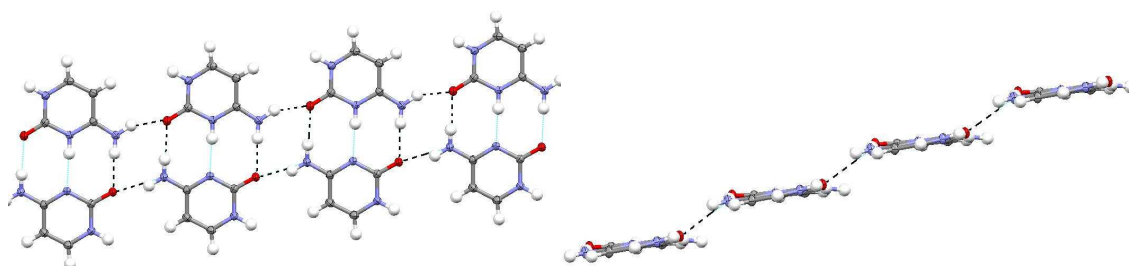


Figure 3.17. Stepped cytosine chain in the cytosine 2-fluorobenzoic acid hemihydrate molecular complex.

The 2-fluorobenzoate molecule is tethered to the neutral side of the cytosine chain via a hydrogen bond of moderate strength with an N...O distance of 2.795(2) Å (Figure 3.18). The benzoic acid molecule lies approximately perpendicular to the cytosine chain. The water plays a significant role in the crystal structure acting as an acceptor for a moderately strong hydrogen bond from a protonated cytosine of N...O distance of 2.698(2) Å on the other side of the chain. The water molecule also provides a link between parallel cytosine chains.

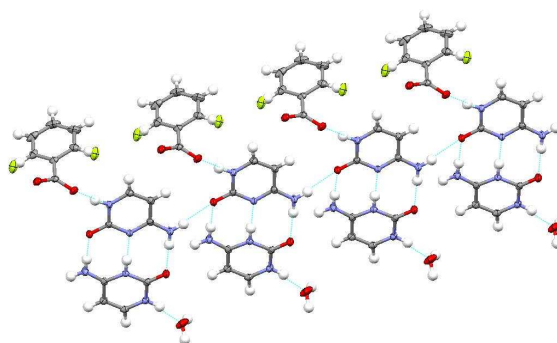


Figure 3.18. 2-fluorobenzoate and water molecules tied to the cytosine chain in the cytosine 2-fluorobenzoic acid hemihydrate molecular complex.

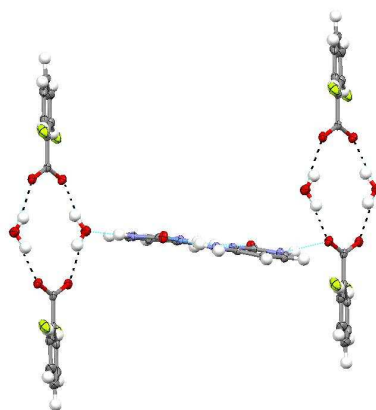


Figure 3.19. Ring shape formed between two water and two 2-fluorobenzoate molecules in the 2-fluorobenzoic acid and cytosine molecular complex.

The water acts as a hydrogen bond donor to two separate 2-fluorobenzoate molecules to form a hydrogen bonded ring, $R_4^4(12)$. The hydrogen bonds are of moderate strength and have O...O distances of 2.731(2) Å and 2.753(2) Å.

From this hydrogen bonded ring, two cytosine, two cytosinium and two 2-fluorobenzoic acid molecules are hydrogen bonded. This ring thus helps interlink different cytosine chains. The four cytosines are part of different cytosine chains running perpendicular to this ring. The 2-fluorobenzoic acid molecules sandwich pairs of parallel cytosine chains and these are aligned head to tail with respect to one another (Figure 3.19). They form O-H...O hydrogen bonds to the carboxylate group of the 2-fluorobenzoate molecules with O...O distances of 2.592(2) Å. In addition, the carbonyl group of the carboxyl group acts as a hydrogen bond acceptor to both an N-H...O and C-H...O hydrogen bond from a parallel cytosinium molecule. The N...O and C...O distances are 2.930(3) Å and 3.111(3) Å, respectively.

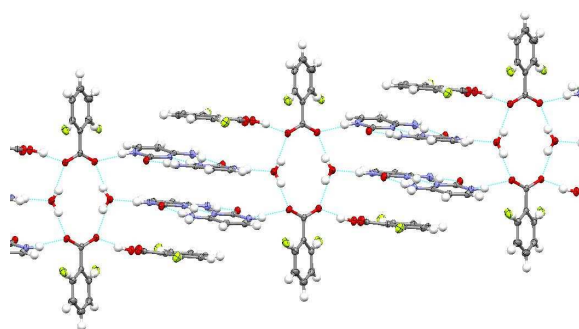


Figure 3.20. Ring connecting layers in the molecular complex of cytosine and 2-fluorobenzoic acid hemihydrate.

When discussing the fluorine interactions in this structure the changes in occupancy present in the two 2-fluorobenzoic acid molecules must be taken into account. For the correct picture to be seen the major component will be described first and then the minor component.

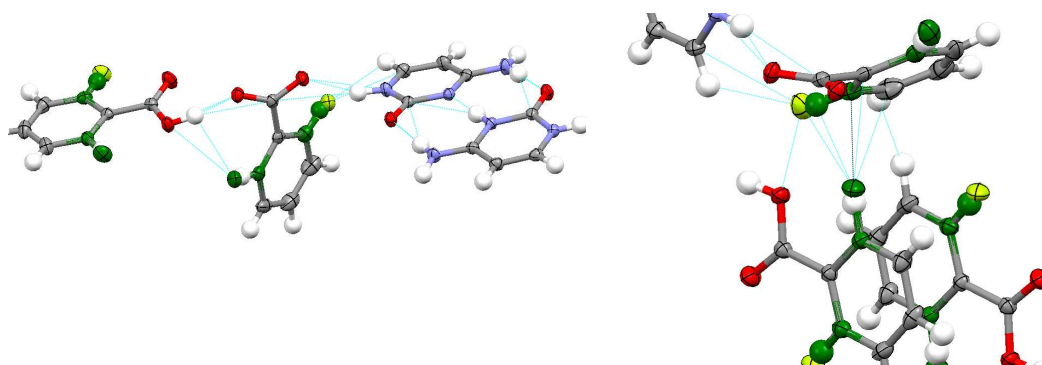


Figure 3.21. Fluorine interaction involving the major component of the disordered 2-fluorobenzoic acid molecule in the molecular complex of cytosine and 2-fluorobenzoic acid hemihydrate.

The major component is 97% occupied and has been colour coded green. When the fluorine on the 2-fluorobenzoate is in the major position the fluorine will be involved in a F...O interaction between the fluorine of the 2-fluorobenzoate and the protonated oxygen atom of the carboxylic acid group of the 2-fluorobenzoic acid. This bonding distance is of length 2.894(2) Å and helps act as a stabilising factor between the two molecules that are perpendicular to one another as seen in Figure 3.21 (left). When the fluorine in the 2-fluorobenzoic acid molecule is positioned in the major position then it will be involved in an interaction between the fluorine atom and a ring carbon that is attached to a hydrogen in the major component. This bonding distance is of length 2.377 Å, which is much shorter than the summed van der Waals radii of carbon and fluorine (3.17 Å). This interaction also acts as a stabilising factor between separate parts of the molecule as seen in Figure 3.21 (right).

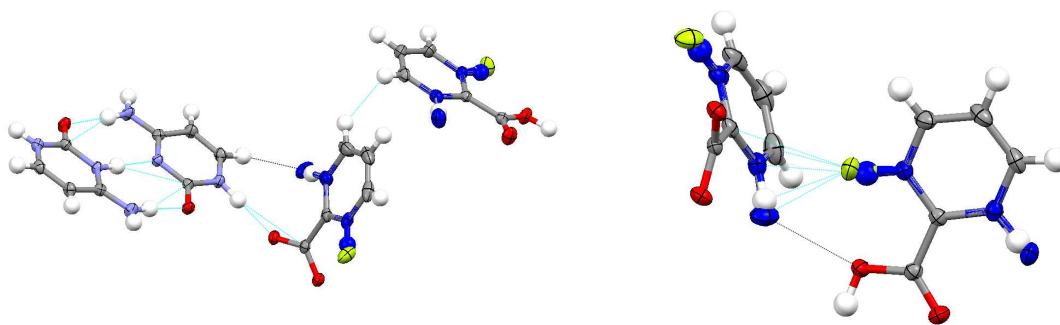


Figure 3.22. Fluorine interaction involving the minor component of the disordered 2-fluorobenzoate (left) and 2-fluorobenzoic acid (right) molecules in the molecular complex of cytosine and 2-fluorobenzoic acid hemihydrate.

When the focus changes to the minor component the interactions obviously change. In this section the minor component has been colour coded blue. When looking at the minor component of the 2-fluorobenzoate there is a moderate strength C-H...F hydrogen bond present in the structure. This is between the fluorine of the 2-fluorobenzoate and a ring C-H group of a nearby cytosine molecule. This bond distance is of moderate strength and is of length 3.024(2) Å. When dealing with the fluorine from the minor component of the 2-fluorobenzoate there is an interaction between the fluorine and the oxygen of the nearby carboxylic acid group of a 2-fluorobenzoic acid molecule. This bond distance is of length 2.645(2) Å, which is very much less than the summed van der Waals radii of oxygen and fluorine (2.99 Å).

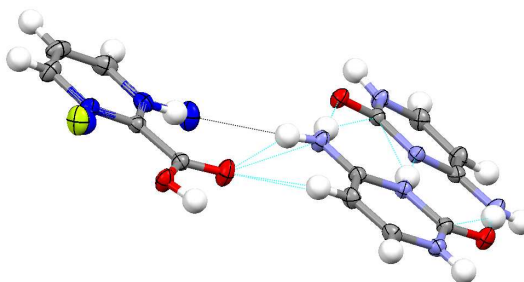


Figure 3.23. F...H-N hydrogen bond present between the fluorine of the 2-fluorobenzoic acid molecule and the amine group of a protonated cytosine molecule in the molecular complex of cytosine and 2-fluorobenzoic acid hemihydrate

The final fluorine interaction of the minor component involves the fluorine of the 2-fluorobenzoic acid molecule. This is involved in a weak strength hydrogen bond with the amine group of a nearby protonated cytosine molecule. This bond is of weak strength and is of length 3.234(2) Å. This interaction acts as an extra stabilising factor between the cytosine and the 2-fluorobenzoic acid molecule tied to the chain as seen in Figure 3.23.

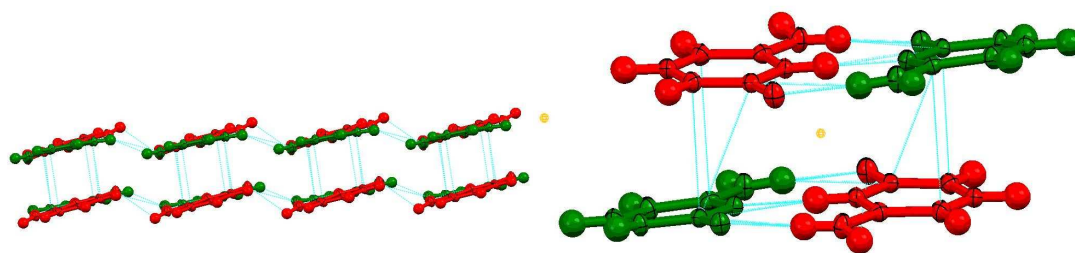


Figure 3.24 Stacking interactions in the molecular complex of cytosine and 2-fluorobenzoic acid hemihydrate. Colour scheme, green represents neutral cytosine and red represents the protonated cytosine molecule.

The stacking interactions here are similar to those seen in the benzoic acid cytosine molecular complex (Section 3.1). The two cytosine base-paired chains stack on top of one another with the charged side of the chain stacking on top of the neutral side of the chain and this is reversed for the other side of the chain (Figure 3.24). These two units are related by an inversion centre. The spacing of the layers is approximately 3.132 Å.

3.3 1:1 Molecular Complex of Cytosinium 3-Fluorobenzoate

Cytosine and 3-fluorobenzoic acid (3FBA) were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a needle shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.5mm by 0.2mm by 0.2mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 3.11.

The molecular complex forms in a 1:1 ratio of cytosine to 3-fluorobenzoic acid. There is 100% hydrogen transfer from the benzoic acid to the cytosine in contrast to the benzoic acid and 2-fluorobenzoic acid complexes. The hydrogen transfers from the carboxyl group of the benzoic acid to the usually unprotonated nitrogen on the cytosine.

In the other related structures the partial hydrogen transfer allows the pseudo-Watson-Crick cytosine-cytosinium base pairing motif to form. This structure cannot have this type of base pairing due to the 100% hydrogen transfer; instead it forms a hydrogen-bonded pair between the cytosinium and the 3-fluorobenzoate molecules through the carboxylate group on the latter (Figure 3.25).

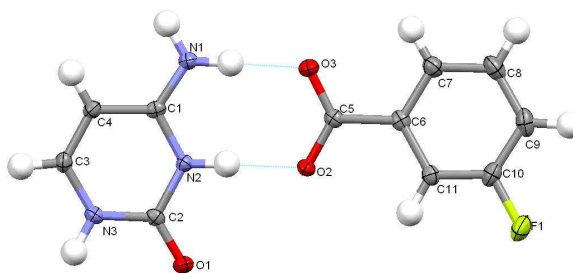


Figure 3.25. The asymmetric unit for the molecular complex of cytosinium 3-fluorobenzoate including atom labelling.

Table 3.5. Hydrogen bond distances of the hydrogen bonded ring in the molecular complex of cytosinium 3-fluorobenzoate.

Hydrogen Bond	N...O Distance
D. N1-H...O3	2.704(2) Å
C. N2-H...O2	2.724(2) Å

This two molecule heterodimer pair is the basic building block for this structure, with hydrogen bonded rings formed of type $R^2_2(8)$. The cytosinium is involved in five hydrogen bonds, two of which comprise the hydrogen-bonded ring as previously mentioned. The remaining three hydrogen bonds are single hydrogen bonds to one 3-fluorobenzoate molecule and two other cytosinium molecules (Figure 3.26). The single hydrogen bond to the benzoic acid has an N...O distance of 2.688(2) Å and is of moderate strength. The two cytosine single hydrogen bonds have an N...O distance of 2.872(2) Å and are equivalent.

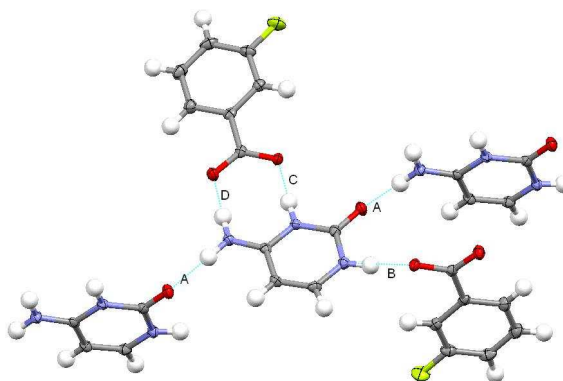


Figure 3.26. Hydrogen bonding involving the heterodimer in the molecular complex of cytosinium 3-fluorobenzoate.

Hydrogen bond B is shorter as it is charge assisted. Bond A creates chains of cytosinium molecules along the *c*-axis. In the chain each alternate cytosinium molecule is perpendicular to the previous one in the chain (Figure 3.27). Parallel chains are connected through hydrogen bonds between cytosinium molecules and 3-fluorobenzoate molecules (Figure 3.29).

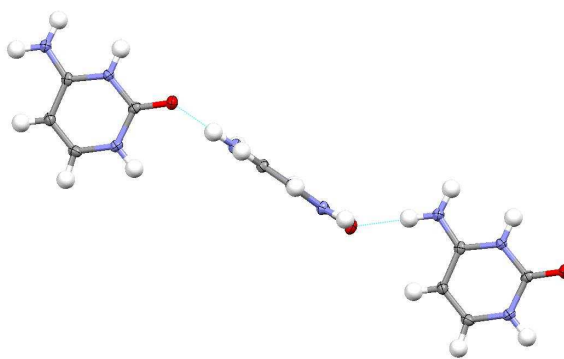


Figure 3.27. Alternating orientation of cytosinium molecules in the cytosinium 3-fluorobenzoate molecular complex.

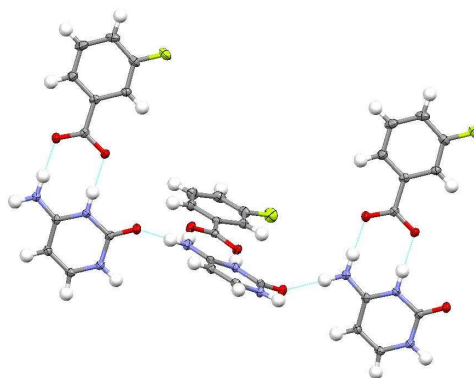


Figure 3.28. Heterodimers connected through hydrogen bonds in the molecular complex of cytosinium 3-fluorobenzoate.

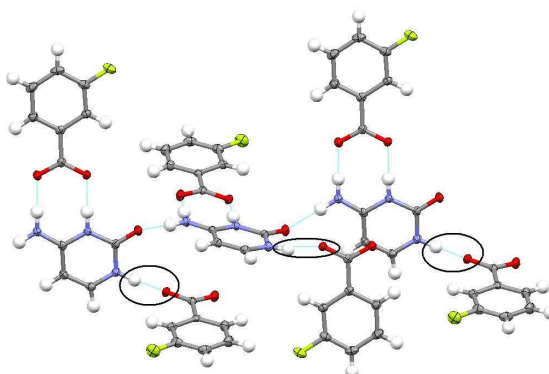


Figure 3.29. Two separate chains (Figure 3.28) linked together via a single hydrogen bond in the cytosinium 3-fluorobenzoate molecular complex.

The fluorine is involved in a weak strength C-H...F DDHHA bifurcated hydrogen bond. Both hydrogen bond donors are CH groups on the cytosinium ring. The major component of this bifurcated hydrogen bond has a C...F distance of 3.478(1) Å and the minor component a C...F distance of 3.529(2) Å. This fluorine hydrogen bond connects molecules that lie in the same plane. The fluorine acts as a link to different chains that would not be possible in the non-fluorinated equivalent.

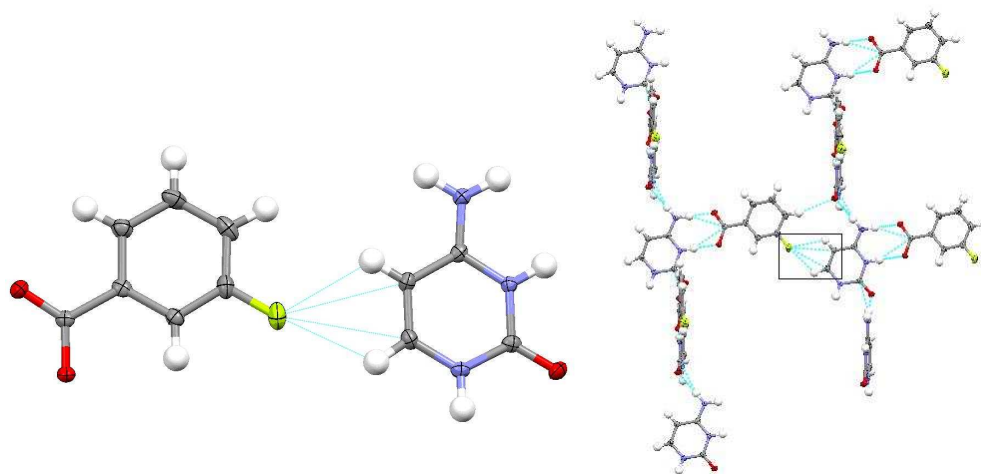


Figure 3.30. Fluorine interactions in the cytosinium 3-fluorobenzoate molecular complex.

There is no base stacking in this structure, however, there are stacking interactions between two 3-fluorobenzoate molecules and between a 3-fluorobenzoate molecule and a cytosinium molecule. There are three stacking interactions with two unique distances. The distances between the green cytosinium molecules and the blue 3-fluorobenzoate molecules are equal (Figure 3.31). The two unique distances are between the two blue 3-fluorobenzoate molecules and the distance between the green cytosinium and the blue 3-fluorobenzoate molecule of approximately 3.136 Å and 3.284 Å, respectively. The structure is not layered and so this is the extent of the stacking interactions (Figure 3.31, right)

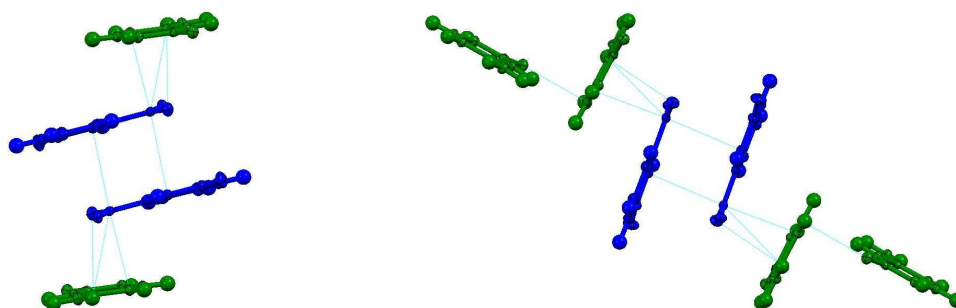


Figure 3.31 Stacking interactions in the cytosinium 3-fluorobenzoate molecular complex. Colour scheme, green represents cytosinium and blue represents the 3-fluorobenzoate molecules.

3.4. 2:1:1 Molecular Complex of Cytosine and 4-Fluorobenzoic acid Hydrate

Cytosine and 4-fluorobenzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.4mm x 0.4mm x 0.3mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 3.11.

The molecular complex forms in a 2:1:1 ratio of cytosine to 4-fluorobenzoic acid to water. Whilst the ratio in the product is represented as 2:1:1, one cytosine has gained a hydrogen from the carboxyl group of the 4-fluorobenzoic acid essentially creating two different molecular species of cytosine, one of which has gained a hydrogen. There are therefore four independent molecules in the asymmetric unit, a neutral cytosine, a protonated cytosinium, a deprotonated 4-fluorobenzoate and a water molecule (Figure 3.32). The proton transfer allows the formation of a pseudo-Watson-Crick base paired dimer between the cytosine and cytosinium molecules.

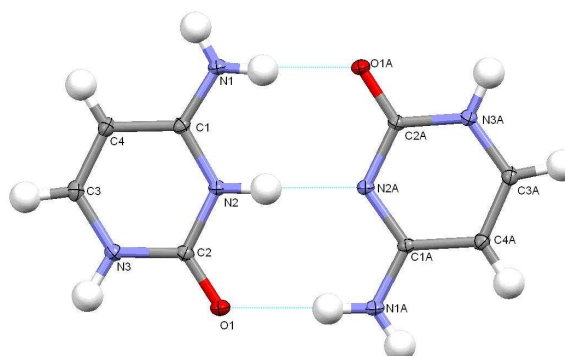


Figure 3.32. The pseudo-Watson-Crick cytosine-cytosinium base pair observed in the molecular complex of 4-fluorobenzoic acid and cytosine hydrate with numbering scheme.

Table 3.6. Hydrogen bond distances of the pseudo-Watson-Crick hydrogen bonding motif in the molecular complex of cytosine and 4-fluorobenzoic acid hydrate.

Hydrogen Bond	D...A Distance
N1A-H---O1	2.949(2) Å
N2A---H-N2	2.836(2) Å
O1A---H-N1	2.729(2) Å

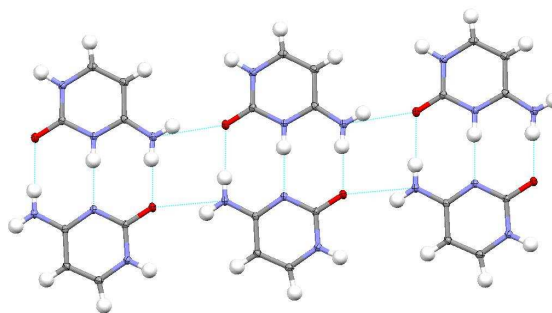


Figure 3.33. Cytosine homo-base paired chain formed in the molecular complex of cytosine and 4-fluorobenzoic acid hydrate with hydrogen bonded squares linking pseudo-Watson-Crick-base pairs along the *b*-axis.

The cytosine dimers form chains along the *b*-axis with two N-H...O hydrogen bonds connecting one base paired unit to another forming an approximately square shaped linker comprised of four hydrogen bonds (Figure 3.33). These hydrogen bonds are of moderate strength and form between the carbonyl groups and the amino groups and have N...O distances of 2.914(1) Å and 2.920(1) Å, on the cytosine and cytosinium sides of the chain, respectively. Hydrogen bonded rings are thus formed of type $R^2_4(8)$. The chains are planar.

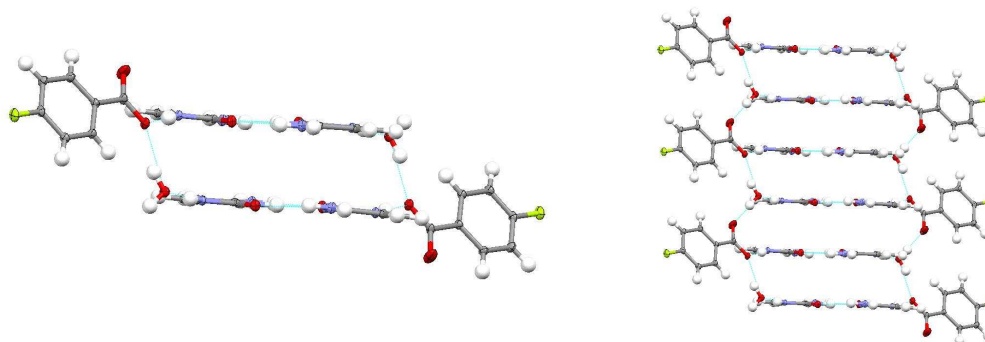


Figure 3.34. Formation of ladder structure in the molecular complex of cytosine and 4-fluorobenzoic acid with the water molecule acting as a spacer.

A combination of hydrogen bonding involving the 4-fluorobenzoate and water molecules connects the layers (Figure 3.34, left). Each oxygen on different 4-fluorobenzoic acid molecules acts as a hydrogen bond acceptor from the water. This creates chains of alternating 4-fluorobenzoate and water molecules. In addition, the longer of the two C-O units of the carboxylate group hydrogen bonds to the cytosine molecule with an N...O distance of 2.691(1) Å. On the other side of the cytosine chain, the cytosinium molecule forms a hydrogen bond to the oxygen of the water molecule with an N...O distance of 2.674(1) Å. This generates a ladder structure with the cytosine chains being the rungs (Figure 3.34, right).

The hydrogen-bonding motif is similar to that seen in the benzoic acid and cytosine molecular complex (Section 3.1) and is displayed in Figure 3.35. The cytosine and benzoic acid molecular complex contains a six molecule unit comprised of two charged cytosinium

molecules, two neutral cytosine molecules and two charged benzoate molecules. The cytosine and 4-fluorobenzoic acid molecular complex has the same composition but with two water molecules included in addition. The water molecule acts as a spacer and this offers extended hydrogen bonding opportunities resulting in the ladder structure which cannot exist in anhydrous complexes.

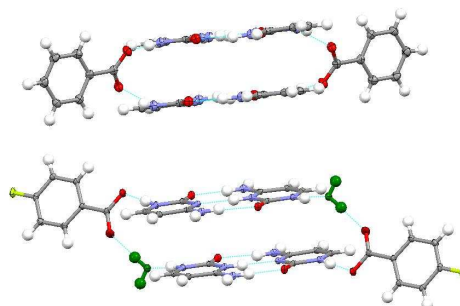


Figure 3.35. Comparison of the six membered block in the molecular complex of cytosine and benzoic acid (top) and the eight membered block in the molecular complex of cytosine and 4-fluorobenzoic acid hydrate (bottom) (the water molecules are in green).

The fluorine atom in this molecular complex is the most exposed to be able to take part in fluorine interactions with separate molecules and chains due to its positioning relative to the carboxylate group. The fluorine is involved in two weak hydrogen bonds with two separate layers. These two hydrogen bonds involve the same hydrogen atom on two distinctly different cytosine molecules, one cytosine and one cytosinium. The hydrogen involved is that on the carbon next to the normally protonated heteratom of the cytosine ring. The hydrogen bond involving the neutral cytosine has a C...F length of 3.595(2) Å. The other hydrogen bond involving the cytosinium molecule has a C...F length of 3.473(1) Å. These interactions act as a link between parallel ladder structures.

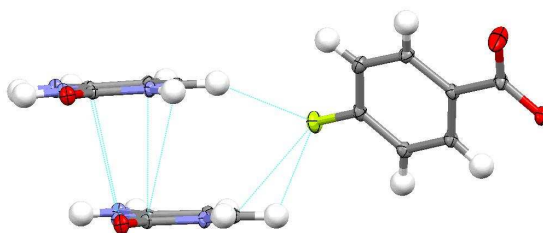


Figure 3.36. Fluorine interactions in the molecular complex of cytosine and 4-fluorobenzoic acid hydrate.

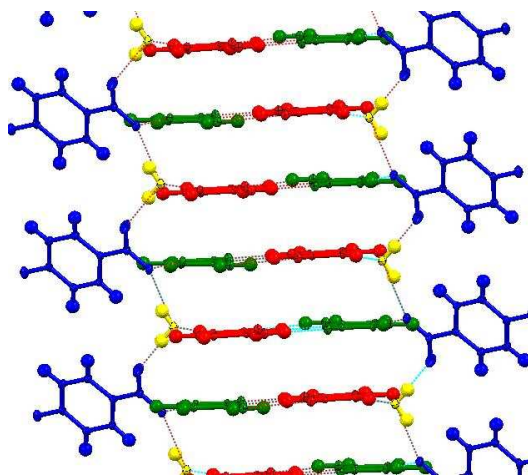


Figure 3.37 Stacking interactions in the molecular complex of cytosine and 4-fluorobenzoic acid hydrate with colour scheme: red representing cytosinium molecules, green representing neutral cytosine molecules, blue representing the 4-fluorobenzoate molecules and yellow representing the water molecules.

The stacking interactions in this structure are base stacking interactions between the separate layers of the cytosine chains. There are two unique stacking interactions between the neutral cytosine chain on top of the charged chain and in turn the charged chain on top of the neutral chain (Figure 3.37). These stacking interactions are found to be at distances of approximately 3.181 and 3.220 Å, respectively.

3.5. Comparison of the Series of Molecular Complexes of Cytosine with Fluorosubstituted Benzoic acids.

Some common structural trends in the cytosine and benzoic acid molecular complexes can be identified. There is hydrogen transfer from a benzoic acid molecule to the unprotonated nitrogen in a cytosine molecule in all four complexes; however, in some cases, both protonated and unprotonated cytosine and benzoic acid molecules co-exist. There is diversity in the crystallisation ratio of the complexes formed and two of the complexes are hydrates.

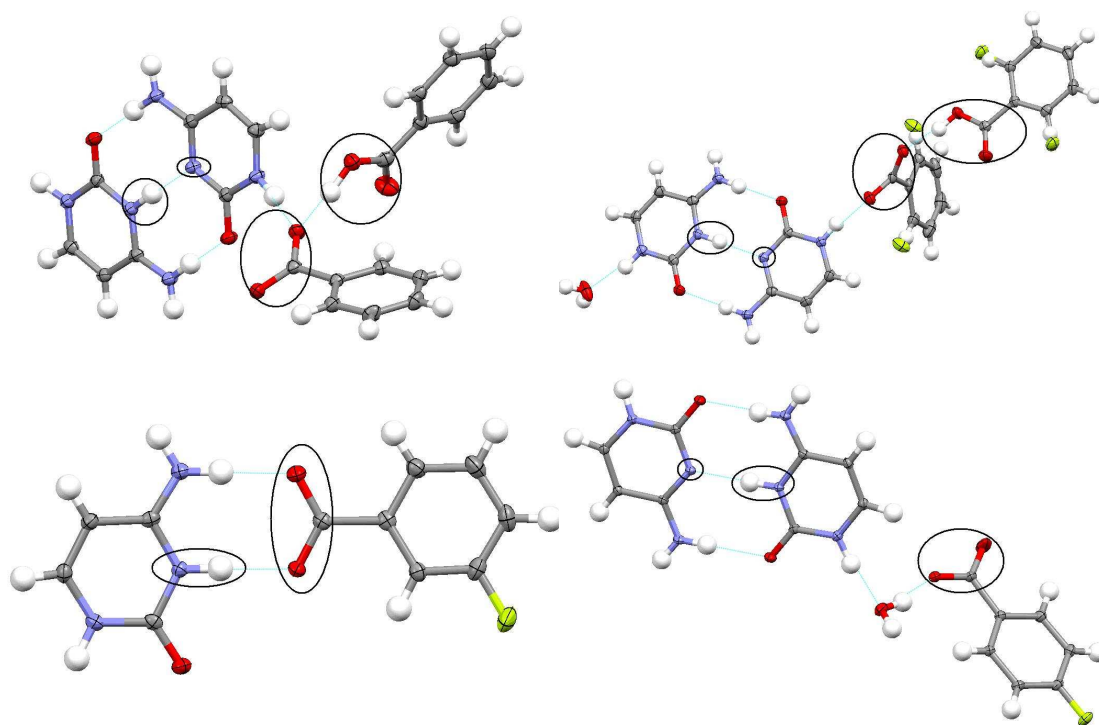


Figure 3.38. The asymmetric units and proton transfer in (a) cytosine and benzoic acid, (b) cytosine and 2-fluorobenzoic acid hydrate (c) cytosinium 3-fluorobenzoate (d) cytosine and 4-fluorobenzoic acid hydrate

3.5.1. Hydrogen Transfer and the Effect on the Crystallisation Ratio of the Product

In all cases there is transfer of a proton from a benzoic acid carboxyl group to the unprotonated nitrogen of the cytosine. However, there is a range of crystallisation ratios observed. The normal method to try and anticipate hydrogen transfer is to take into account the ΔpK_a values between the acid and the base components.

Table 3.7 The ΔpK_a values, level of hydrogen transfer and the ratio of the components in the four molecular complexes

Molecular complex	ΔpK_a	Hydrogen transfer with respect to cytosine molecules	Crystallisation Ratio
Cytosine and benzoic acid	0.396	50%	1:1
Cytosine and 2-fluorobenzoic acid hydrate	1.33	50%	1:1:1
Cytosinium 3-fluorobenzoate	0.74	100%	1:1
Cytosine and 4-fluorobenzoic acid hydrate	0.45	50%	2:1:1

All of the ΔpK_a values sit in the continuum between 0 and 3 where it is difficult to predict the level of hydrogen transfer. The only complex to show 100% hydrogen transfer with

respect to the cytosine molecule is the cytosinium 3-fluorobenzoate molecular complex, but this interestingly does not show the highest ΔpK_a value. However the reason for the different crystallisation ratio in the cytosine 2-fluorobenzoic acid molecular complex (with the highest ΔpK_a value) is likely to be due to the presence of water in this complex.

Table 3.8 Crystallisation ratio, number of molecular species and presence of water for the four structures

Co-molecule	Ratio not including hydrogen transfer (Cyto:BA:H ₂ O)	Number of molecular species	Hydrate
Benzoic Acid	2:2	4	No
2-Fluorobenzoic acid hydrate	2:2:1	5	Yes
3-Fluorobenzoic acid	1:1	2	No
4-Fluorobenzoic acid hydrate	2:1:1	4	Yes

The benzoic acid and cytosine molecular complex contains equal proportions of the protonated and unprotonated forms of both components. So whilst the ratio in the product is represented as 1:1, one cytosine has gained a hydrogen from the carboxyl group of the benzoic acid essentially creating four different molecular species, two of which carry a charge; a neutral benzoic acid and a deprotonated benzoate ion, and a cytosinium and a neutral cytosine. The 2-fluorobenzoic acid cytosine molecular complex behaves in much the same manner as the benzoic acid complex, however a water is incorporated into the structure. So the ratio in the product is represented as 2:2:1, but one cytosine has gained a hydrogen from the carboxyl group of the benzoic acid essentially creating five different molecular species. The 3-fluorobenzoic acid cytosine molecular complex is the most simple system and does not follow the general trend, as it has 100% hydrogen transfer from one cytosine to one 3-fluorobenzoic acid. There are no neutral molecules present in the structure and so the ratio of this complex is 1:1. The 4-fluorobenzoic acid cytosine molecular complex lies between these two extremes; there is once again hydrogen transfer from the benzoic acid molecule to a cytosine. However, there is no neutral benzoic acid molecule incorporated within the complex. The ratio is represented as 2:1:1; one cytosine has gained a proton from the benzoic acid essentially creating four different molecular species, two of which carry a charge; a deprotonated benzoic acid, a cytosinium and a neutral cytosine with one water molecule also present. This is summarised in Table 3.8.

3.5.2. Structural Motifs of the Cytosine Molecules

The 3-fluorobenzoic acid structure has total transfer of a hydrogen from the benzoic acid to the cytosine. This means that there is no possibility of the pseudo-Watson-Crick forming and instead a hydrogen-bonded ring is formed between the 3-fluorobenzoic acid and the cytosine (Figure 3.25).

The other three structures contain two cytosines, one protonated and one neutral. As a result of this each structure is capable of forming the pseudo-Watson-Crick base-pair motif and all three structures do (Figure 3.39).

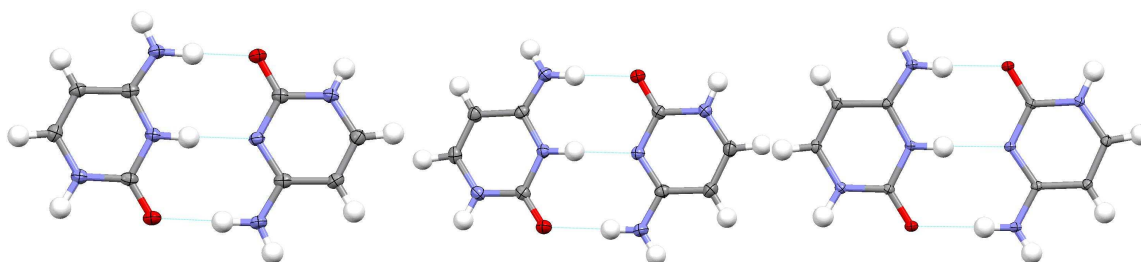


Figure 3.39. Pseudo-Watson-Crick bonding motif found in the molecular complexes of left, cytosine and benzoic acid, middle, cytosine and 2-fluorobenzoic acid hydrate and right, cytosine and 4-fluorobenzoic acid hydrate.

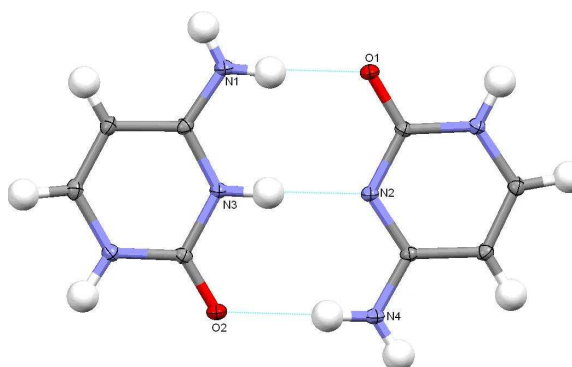


Figure 3.40. Numbering scheme for the pseudo Watson Crick base pair applicable to all three structures.

The pseudo-Watson-Crick hydrogen bonding motifs found in each of these complexes contain three hydrogen bonds. However, there are noticeable differences found in the three. The structures containing benzoic acid and 4-fluorobenzoic acid have very similar hydrogen bonding lengths and this could be expected as similar chains of these units are formed in the same fashion. The chains in each are planar, however, this cannot be said for the structure with 2-fluorobenzoic acid. Here all the bonds are very similar in length and the change may be due to the change in the construction of the chains with a stepped chain being adopted. However, the bonding between separate base pairs is also lengthened due to the change in chain formation.

Table 3.9 Comparison of the three pseudo-Watson-Crick bonding motifs seen in the molecular structures of cytosine with benzoic acid and 2- and 4-fluorobenzoic acid

	Pseudo Watson Crick		
	Benzoic acid	2-Fluorobenzoic acid	4-Fluorobenzoic acid
O1-N1	2.771(2)	2.836(2)	2.730(1)
N2-N3	2.844(2)	2.847(2)	2.836(1)
O2-N4	2.914(2)	2.897(2)	2.949(1)

In these structures each cytosine base pair is linked to another base pair through hydrogen bonding. The main difference is whilst the benzoic acid and the 4-fluorobenzoic acid structures have two hydrogen bonds linking each separate base pair that form a square shape, the link present in the 2-fluorobenzoic acid structure contains a two hydrogen bonded link between base pairs that form a diamond shape. The BA and 4-fluorobenzoic acid structures have a planar chain whilst the 2-fluorobenzoic acid structure contains a more staggered chain.

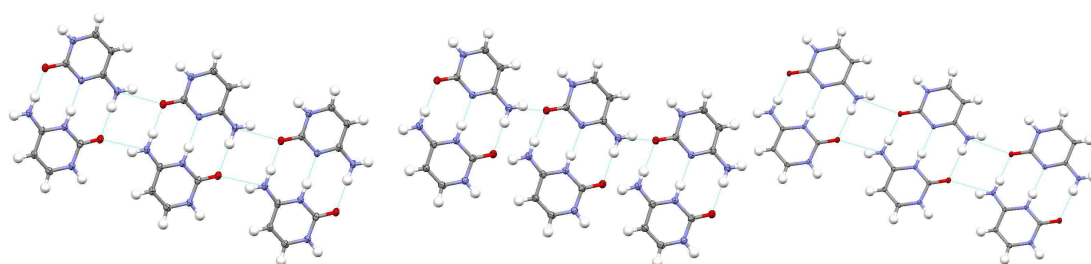


Figure 3.41. Cytosine homo-base paired chain found in the molecular complexes of left, cytosine and benzoic acid, middle, cytosine and 2-fluorobenzoic acid hydrate and right, cytosine and 4-fluorobenzoic acid hydrate.

Table 3.10. Hydrogen bond distances connecting the pseudo-Watson-Crick base pairs into cytosine chains.

	Link Between Cytosine base pairs in chain		
	Benzoic acid	2-Fluorobenzoic acid	4-Fluorobenzoic acid
N1-O1	2.868(2) Å	3.258(2) Å	2.913(1) Å
N2-O2	2.876(2) Å	3.022(2) Å	2.920(1) Å

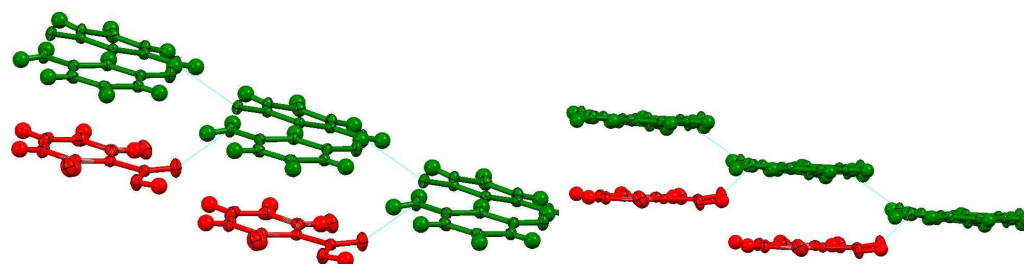


Figure 3.42. Staggered chain in the molecular complex of cytosine and 2-fluorobenzoic acid hydrate with colour scheme: green for cytosine molecules and red for 2-fluorobenzoic acid molecules.

The staggered cytosine chain in the cytosine 2-fluorobenzoic acid molecular complex is not straight like others seen in the series and it contains a slight shift to the right as the

connecting hydrogen bonded units are diamond shaped with much weaker hydrogen bonds involved. There are also some interactions that occur in the 2-fluorobenzoic acid molecular complex that are not present in the other two related structures. This interaction is a moderate strength hydrogen bond of a type N-H...O bond between the amine group of the cytosinium molecule in the chains and the oxygen atom of a carboxyl group of a nearby neutral 2-fluorobenzoic acid molecule. This extra interaction also explains the step in the chains as the positioning of the 2-fluorobenzoic acid molecule cannot be accommodated otherwise (Figure 3.42). This also accounts for the increased distance between base pairs in the chain as seen in Table 3.10.

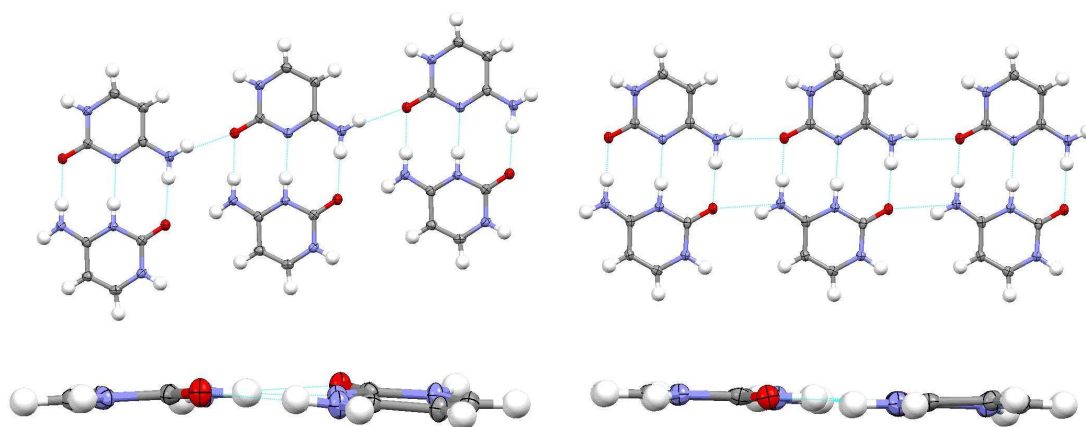


Figure 3.43. Comparing chains and twisting of the base pairs in the molecular complex of cytosine and 2-fluorobenzoic acid (left) and cytosine and benzoic acid (right).

The pseudo-Watson-Crick unit in the 2-fluorobenzoic acid molecular complex is also twisted (Figure 3.43). This, in combination with the step in the cytosine chain, accounts for the main differences in the distances observed in the link hydrogen bonds between base pairs in the 2-fluorobenzoic acid and cytosine molecular complex when compared with the cytosine with benzoic acid and 4-fluorobenzoic acid with cytosine molecular complexes.

3.5.3 Structural Motifs Involving the Benzoic Acid Molecules

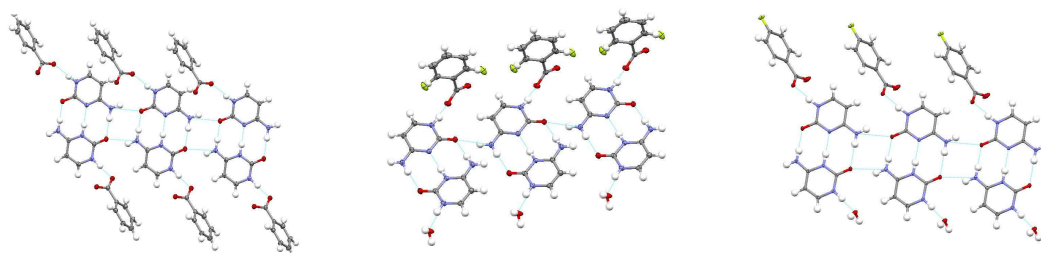


Figure 3.44. Changes in the molecules tied to either side of the homo-base paired chain found in the molecular complexes of left, cytosine and benzoic acid, middle, cytosine and 2-fluorobenzoic acid hydrate and right, cytosine and 4-fluorobenzoic acid hydrate

There are further related hydrogen bonding patterns and structural build up in all three structures involving the benzoic acid and water molecules. In the benzoic acid cytosine molecular complex, there is a benzoic acid molecule tied to the cytosine chain on either side via a single hydrogen bond of a type $N-H\cdots O$. In the other two related structures on the side with the cytosinium molecule, there is a water molecule tied to the chain via a single hydrogen bond and on the other side there is a 2-fluorobenzoic acid molecule tied to the chain with both of these bonds being of type $N-H\cdots O$.

In all of the structures, the benzoic acid molecules link several cytosine chains together. In the benzoic acid structure the benzoic acid molecule links two cytosine chains together via single hydrogen bonds to each chain involving the carboxylate group of the benzoic acid. The benzoic acid molecule hydrogen bonds to one cytosinium molecule in one chain and one neutral cytosine in the other chain and this completes a six molecule unit. However, in the 4-fluorobenzoic acid molecular complex, the water molecule acts as a spacer and sits in the middle of one of the bonds that the 4-fluorobenzoic acid molecule makes with one of the chains. This increases the bonding opportunities and results in a ladder structure being built via hydrogen bonding rather than weak interactions linking the six molecule units of the benzoic acid and cytosine molecular complex. These two structures are very similar despite the introduction of the water molecule.

The 2-fluorobenzoic acid molecular complex is related by the fact that cytosine chains form but differs significantly in the role of the 2-fluorobenzoic acid co-molecules. However, the 2-fluorobenzoic acid molecules are still tethered to the neutral side of the cytosine chain and the water molecules are tethered to the cytosinium side of the chain in a similar manner to the 4-fluorobenzoic acid molecular complex.

3.5.4. Base-pair stacking and the influence of water

In the benzoic acid cytosine molecular complex, different chains are linked together via the carboxylate group of the benzoic acid. The neutral benzoic acid hydrogen bonds to one of the oxygen atoms of the carboxylate group to stabilise the charge. Due to the benzoic acid only linking two chains there are only weak interactions linking separate six molecule units such as π - π stacking (Figure 3.45).

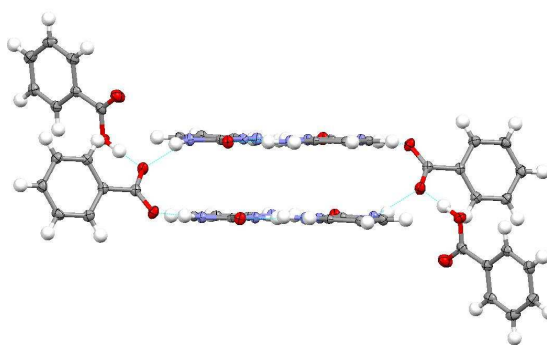


Figure 3.45. The six membered block found in the molecular complex of cytosine and benzoic acid with the neutral benzoic acid tied to the block via a hydrogen bond.

The water incorporated into two of the structures helps to create a link between multiple chains. It is involved in both the 2-fluorobenzoic acid and the 4-fluorobenzoic acid structures. In the 2-fluorobenzoic acid molecular complex, the water helps to form a hydrogen bonded ring structure and from this unit many layers are linked together (Figure 3.46). In the 4-fluorobenzoic acid structure the water plays a more simplistic role combining with the benzoic acid to build up a backbone from which the structure can build, similar to that seen in the benzoic acid complex (Figure 3.47). The water acts as a hydrogen bond acceptor for a hydrogen bond from a cytosinium molecule in a chain and then hydrogen bonds to two separate benzoic acids as a donor.

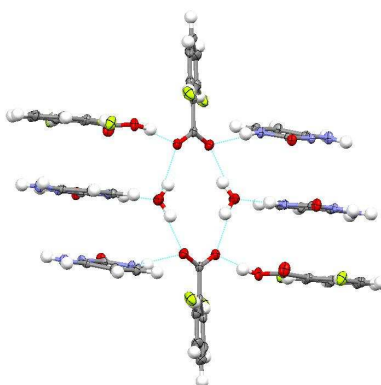


Figure 3.46. Ring shape helping to link several layers together via hydrogen bonding in the 2-fluorobenzoic acid and cytosine molecular complex.

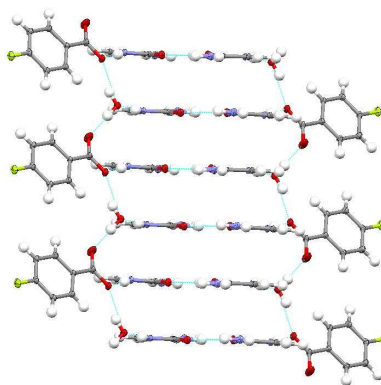


Figure 3.47. Formation of ladder structure in the molecular complex of cytosine and 4-fluorobenzoic acid with the water molecule acting as a spacer.

The 2-fluorobenzoic acid molecular complex shows a shorter stacking distance (3.132 Å) and this could be a consequence of the differing construction of the cytosine chains. However, the base stacking distances are broadly similar in all the complexes and the stacking pattern is consistently neutral cytosine molecules stacked upon the charged cytosinium molecules. The water molecules acting as spacers do not significantly affect the base stacking. The stacking interactions all fall within the range 3.1 to 3.3 Å.

3.5.5 The Anomalous Structure of the Cytosine with 3-Fluorobenzoic Acid Molecular Complex

There is one structure that does not contain any similarities to the other structures and that is the 3-fluorobenzoic acid with cytosine molecular complex. There is also no mixed protonation of the cytosine molecule in this structure; instead there is 100% transfer of a proton from the 3-fluorobenzoic acid to the cytosine. This structure forms a hydrogen bonded ring directly between the cytosinium and the 3-fluorobenzoate molecules. This difference is down to the total hydrogen transfer, meaning there is not a neutral cytosine and a charged cytosine present in the structure which makes the pseudo-Watson-Crick motif impossible to form. There are single hydrogen bonds from one cytosine to another but no homo base-pairing present in the structure.

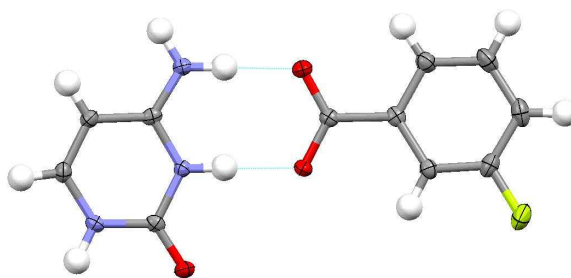


Figure 3.48. Formation of the hydrogen bonded ring in the molecular complex of 3-fluorobenzoic acid and cytosine.

3.5.6 Stacking Interactions

The stacking interactions in the three related structures are very similar. In each of these structures there is base-stacking with the homo-base paired chains stacking on top of each other. In all of the structures the side of the chain containing the neutral cytosine stacks on top of the charged side of another chain. As can be seen from Figure 3.49 none of the structures contain perfectly overlaid chains; they are staggered with respect to one another. This seems to be a common occurrence for structures that contain the homo-base paired chain.

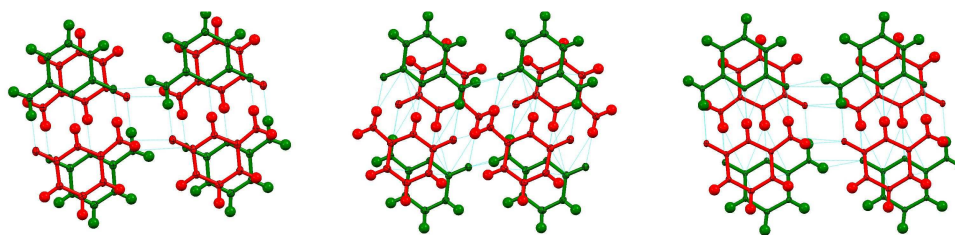


Figure 3.49 Base stacking interactions seen in the molecular complexes of (a) cytosine and benzoic acid, (b) cytosine and 2-fluorobenzoic acid hydrate and (c) cytosine and 4-fluorobenzoic acid hydrate

Table 3.11. Summary of crystallographic data for the molecular complexes of cytosine with benzoic acid and 2-,3- and 4-fluorobenzoic acid.

Compound	Cytosine and benzoic acid	Cytosine and 2-fluorobenzoic acid hydrate	Cytosinium and 3-fluorobenzoate	Cytosine and 4-fluorobenzoic acid
Formula	C22 H22 N6 O6	C22 N6 F2 O7 H22	C11 N3 O3 F1 H10	C7 N6 F1 O5 H17
Crystallisation Conditions	Methanol, Room Temperature	Methanol, Room temperature	Methanol, Room Temperature	Methanol, Room temperature
Molecular weight / g mol⁻¹	466.46	520.44	251.21	284.24
Temperature (K)	100 K	100 K	100 K	100 K
Space Group	P2 ₁ /c	P-1	P2 ₁ /c	P-1
<i>a</i> (Å)	22.330(6)	7.3190(4)	10.0540(2)	6.7093(5)
<i>b</i> (Å)	7.4333(2)	12.6542(8)	8.3894(2)	7.4670(6)
<i>c</i> (Å)	12.8085(3)	13.1862(7)	13.8080(2)	17.6894(1)
α (°)	90	74.872(2)	90	87.774(4)
β (°)	92.461(1)	77.657(2)	110.284(1)	84.370(4)
γ (°)	90	86.658(2)	90	67.464(4)
Volume (Å³)	2124.35(9)	1151.68(8)	1092.44(4)	814.59(15)
Z	4	2	4	2
θ range (°)	1-27.485	3.2-27.5	2.2-27.5	1.2-35.2
Completeness	98.10%	98.70	99.80	95.00
Reflections Collected	28920	14252	19258	56320
Independent	4887	5232	2495	6908
Refln (obs.I>2 σ (I))	3174	4003	1693	5425
R_{int}	0.0692	0.0272	0.634	0.0367
Parameters	395	422	203	312
GooF on F²	1.020	1.084	1.022	1.048
R₁ (Observed)	0.049	0.045	0.043	0.045
R₁ (all)	0.093	0.061	0.083	0.060
wR₂ (all)	0.1059	0.133	0.098	0.130

3.6. Molecular Complexes of 5-Fluorocytosine with Benzoic acid and the Series of Fluorobenzoic Acids

The structures present in this section are the 5-fluorocytosine structures with the co-molecules used in the previous four structures. These structures contain little in the way of structural similarities with the non-fluorinated equivalent with different hydrogen bonded motifs adopted primarily driven by a change in the hydrogen transfer present in these structures.

3.6.1. 1:1 Molecular Complex of 5-Fluorocytosine and Benzoic acid

5-Fluorocytosine and benzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.3mm x 0.3mm x 0.3mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 3.14 at the end of this Section.

The molecular complex forms in a 1:1 ratio of 5-fluorocytosine to benzoic acid (Figure 3.50). In this structure there is no hydrogen transfer and it is anhydrous. This means the pseudo-Watson-Crick base pair cannot be formed and the type A bonding motif predominates (Section 1.10). The cytosine molecules form an extended chain in the *ab* direction through type A and type B hydrogen bonding motifs, which alternate along the chain (Figure 3.51). There is an inversion centre in the middle of each hydrogen-bonded unit. The hydrogen bonds that make up the type A base pair are both of length 2.963(2) Å and the hydrogen bonds making up the type B base pairing motif are both of length 2.794(2) Å, both of moderate strength.

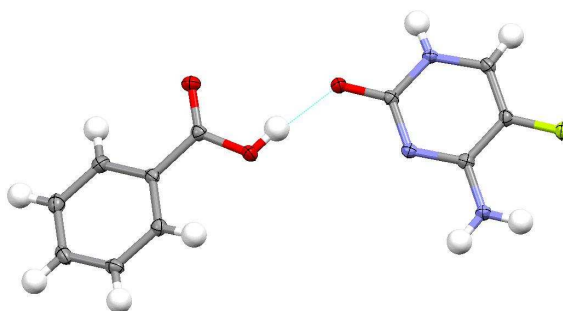


Figure 3.50. The asymmetric unit of the molecular complex of 5-fluorocytosine and benzoic acid.

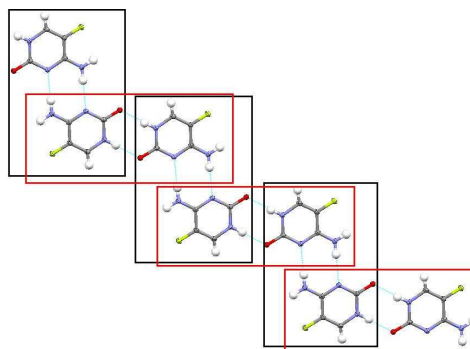


Figure 3.51 Alternating type A (black box) and type B (red box) hydrogen bonding motifs creating an extended 5-fluorocytosine chain in the molecular complex of 5-fluorocytosine and benzoic acid.

The benzoic acid molecules are connected to the main cytosine chain through three hydrogen bonds, two involving the hydroxyl oxygen of the carboxylic acid group: in the first, the benzoic acid acts as a proton donor to the carbonyl group on the cytosine (black in Figure 3.52); in the second, it acts as a proton acceptor from a neighbouring amine group of a second cytosine molecule (red in Figure 3.52). This second cytosine molecule also acts as a hydrogen bond acceptor through the fluorine atom, forming a C-H...F hydrogen bond to an aromatic hydrogen on the benzoic acid molecule (green box, Figure 3.52). There is an additional C-H...O hydrogen bond from a third cytosine molecule to the carbonyl oxygen of the carboxylic acid group (blue in Figure 3.52).

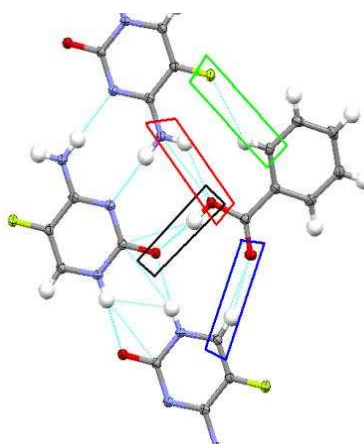


Figure 3.52. Hydrogen bonds found in the molecular complex of 5-fluorocytosine and benzoic acid.

Table 3.12 Hydrogen bonds in the molecular complex of 5-fluorocytosine and benzoic acid with hydrogen bond distances colour coded with Figure 3.52.

Hydrogen bond tethering the benzoic acid to the cytosine chain (atom of benzoic acid first)	D...A length of the hydrogen bond (Å)
1. O-H...O	2.671(2) Å
2. O...H-N	2.900(2) Å
3. C-H...F	3.253(2) Å
4. O...H-C	3.064(2) Å

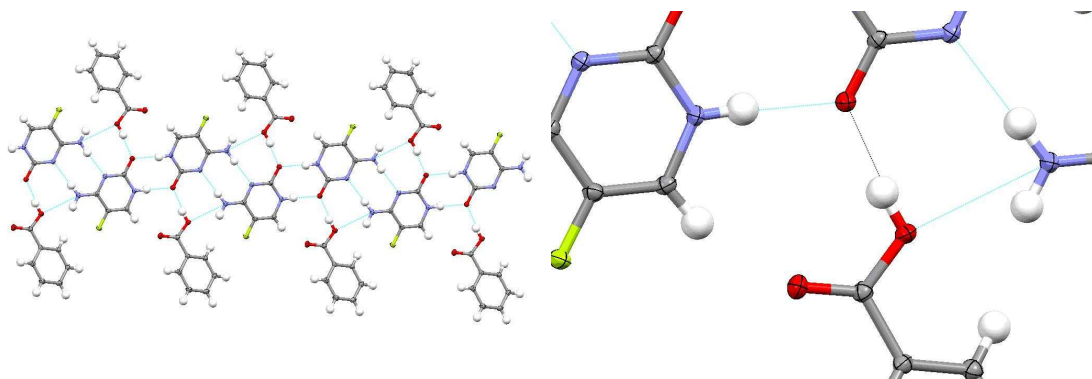


Figure 3.53. Extended cytosine chain with benzoic acid molecules tied to the chain (left) via a moderate strength [130] O-H...O hydrogen bond in the molecular complex of 5-fluorocytosine and benzoic acid.

The shortest of these hydrogen bonds is between the hydroxyl oxygen of the benzoic acid and the oxygen of the cytosine with an O...O distance of 2.671(2) Å and is of moderate strength. The benzoic acid molecule itself is not planar with the carboxylic acid group twisted out of the plane of the benzene ring by approximately 24°. The carboxylic acid group of the benzoic acid molecule is approximately co-planar with the cytosine chain (Figure 3.54).

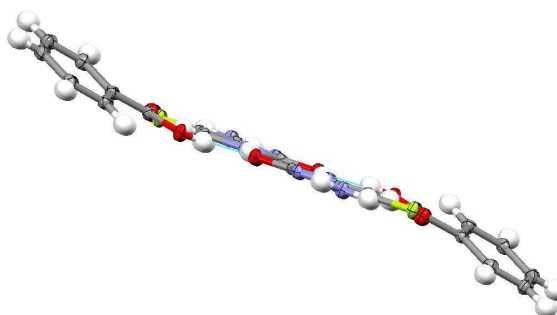


Figure 3.54. The planar nature of the cytosine chains in the molecular complex of 5-fluorocytosine and benzoic acid.

The fluorine on the 5-fluorocytosine molecule plays a connecting role between cytosine chains forming a hydrogen bond with the aromatic hydrogen in the four position of the benzoic acid on a neighbouring chain which is tilted at approximately 65° to the original chain (Figure 3.55). This C-H...F hydrogen bond is of C...F distance of 3.321(2) Å. Again,

the fluorine acts as a link between separate chains but does not interrupt the strong base pairing (Figure 3.56).

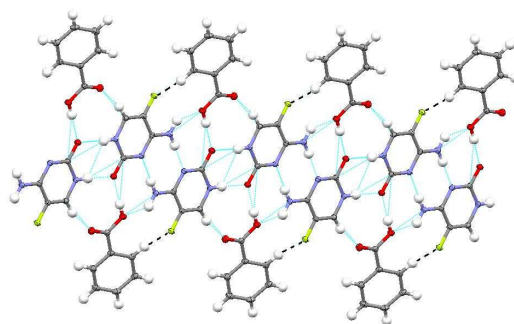


Figure 3.55. Fluorine interactions to the cytosine chain in the molecular complex of 5-fluorocytosine and benzoic acid.

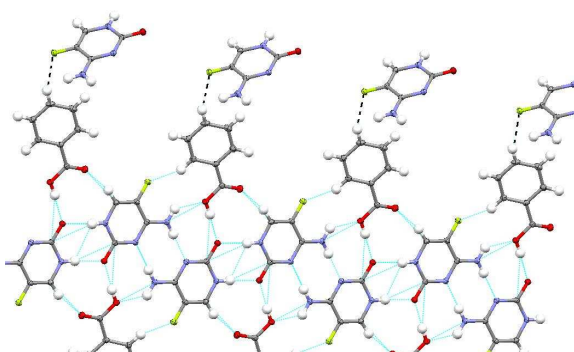


Figure 3.56. Fluorine interactions between separate chains in the molecular complex of 5-fluorocytosine and benzoic acid.

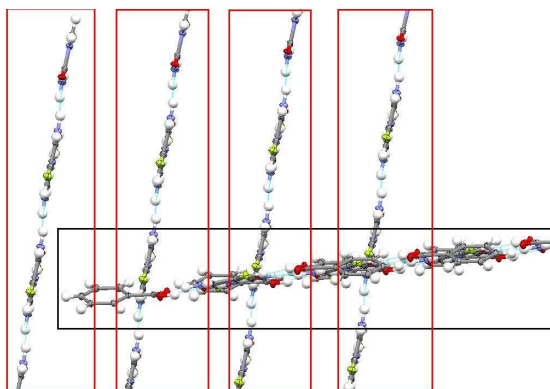


Figure 3.57. Orientation of chains (red) that are connected to a second chain (black) in the molecular complex of 5-fluorocytosine and benzoic acid.

There are several stacking interactions in this structure. The first stacking interaction is between two benzoic acid molecules (Figure 3.58) and this is staggered in nature with the two molecules not overlaying perfectly. The stacking interaction is approximately 3.178 Å in distance. The next stacking interaction is between a 5-fluorocytosine molecule and a benzoic acid molecule stacked immediately on top of it. These molecules are also staggered in relation to one another. The stacking interaction distance is approximately

3.199 Å. The third stacking interaction is present, between a 5-fluorocytosine molecule and the benzoic acid molecule stacked immediately below this. This stacking interaction distance is approximately 3.189 Å.

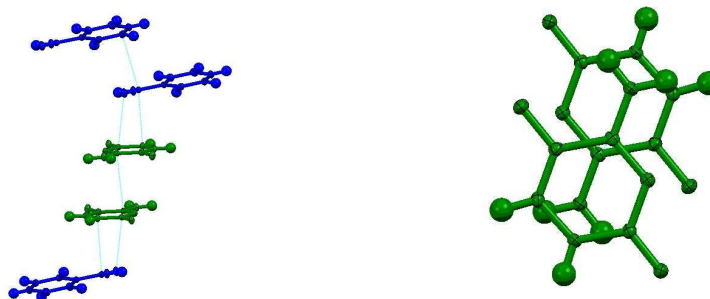


Figure 3.58. Stacking interactions (left) and close up of base stacking interaction (right) found in the molecular complex of 5-fluorocytosine and benzoic acid.

The final stacking interaction is a base stacking interaction between two 5-fluorocytosine molecules. The 5-fluorocytosine molecules stacked on top of each other do not line up as one is flipped by 180 degrees relative to the other as seen in Figure 3.58, right. This means that the bases are staggered and do not overlay exactly. The stacking interaction found here was of approximate distance 3.268 Å.

3.6.2. 1:1 Molecular Complex of 5-Fluorocytosine and 2-Fluorobenzoic acid

5-Fluorocytosine and 2-fluorobenzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.3mm x 0.3mm x 0.3mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 3.14.

The molecular complex forms in a 1:1 ratio of 5-fluorocytosine to benzoic acid. Again, there is no hydrogen transfer and the two pseudo-Hoogsteen motifs predominant in the same manner as for the benzoic acid 5-fluorocytosine molecular complex. This creates the alternating type A and type B base paired chains along the *ab* direction.

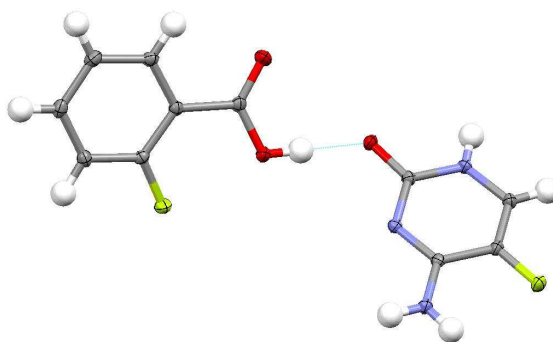


Figure 3.59. Asymmetric unit of the molecular complex of 5-fluorocytosine and 2-fluorobenzoic acid.

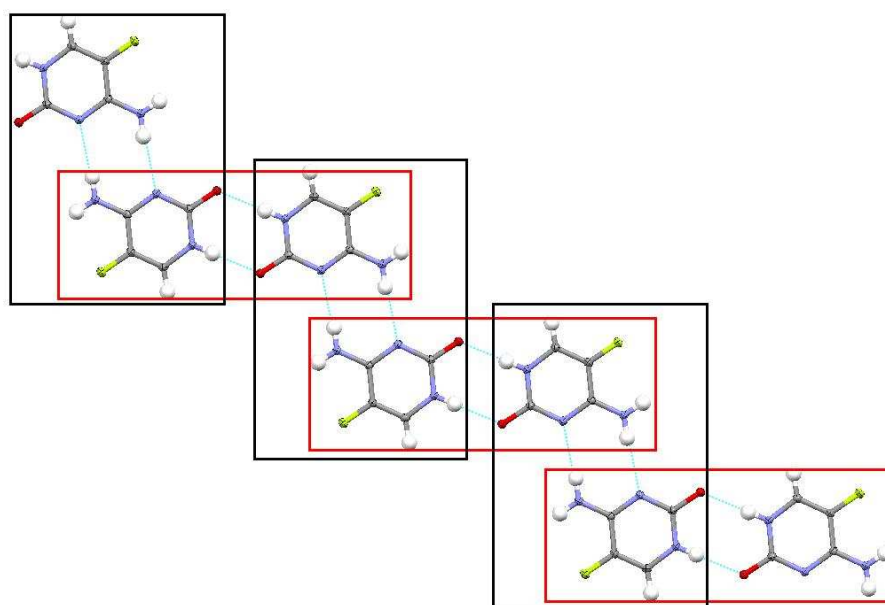


Figure 3.60. The alternating type A (red) and type B (black) base pairing in the molecular complex of 5-fluorocytosine with 2-fluorobenzoic acid.

The 2-fluorobenzoic acid molecules play the same role as before, but this time the O-H...O hydrogen bond between the carboxylic acid group of the 2-fluorobenzoic acid and the cytosine oxygen atom is much shorter with an O...O distance of 2.556 (1) Å (blue in Figure 3.61). Consequently the weak C-H...O hydrogen bond to a neighbouring 5-fluorocytosine molecule in the chain is longer at 3.318 (2) Å (orange in Figure 3.61) and a bifurcated hydrogen bond from the amine group on the 5-fluorocytosine is formed to the fluorine atom and the hydroxyl oxygen of the carboxylic acid group (black in Figure 3.61). This creates a moderate strength N-H...F hydrogen bond with an N...F length of 3.150 (1) Å (green in Figure 3.61). This time, the carboxylic acid group is not co-planar to the cytosine chain and this allows a second C-H...O hydrogen bond to be formed to the carbonyl oxygen on the 2-fluorobenzoic acid molecule of a neighbouring chain. This second hydrogen bond is of moderate strength and has a C...O length of 2.925 (1) Å. The carboxylic acid group is twisted by approximately 50° to the plane of the benzene ring and

this allows the formation of an F...F interaction to a 5-fluorocytosine in the same chain of length 2.721 (2) Å (red in Figure 3.61). It is the combination of two fluorine interactions that causes the 2-fluorobenzoic acid molecule to sit at the angle it does relative to the 5-fluorocytosine chain. This means that the 2-fluorobenzoic acid molecule is tied to the chain via three hydrogen bonds and one F...F interaction (Figure 3.62).

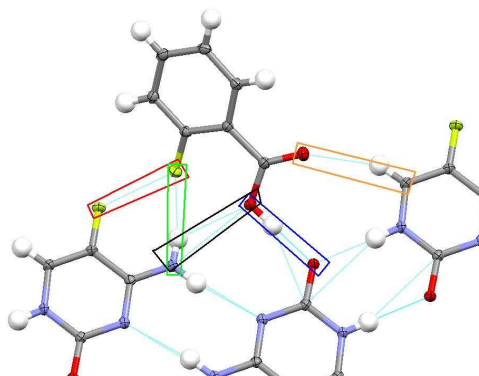


Figure 3.61. Hydrogen bonds connecting the benzoic acid to the 5-fluorocytosine chain in the molecular complex of 5-fluorocytosine and 2-fluorobenzoic acid.

Table 3.13. Hydrogen bonds in the molecular complex of 5-fluorocytosine and 2-fluorobenzoic acid with bond distances colour coded with Figure 3.61.

Hydrogen bond tethering the benzoic acid to the cytosine chain (atom of benzoic acid first)	Length of the bond (Å)
1. F...F	2.721 (2) Å
2. N-H...F	3.150 (1) Å
3. O-H...O	2.556 (1) Å
4. O...H-C	3.318 (2) Å
5. N-H...O	2.925 (1) Å

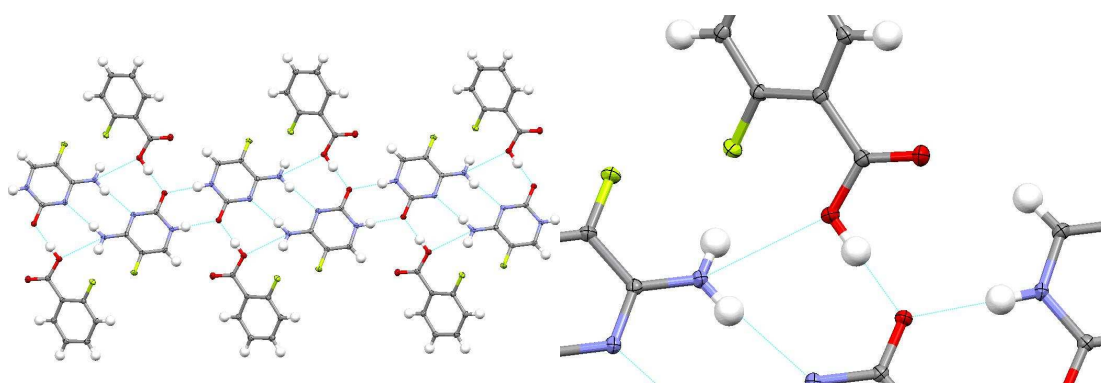


Figure 3.62. Extended chain (left) and the main connection between the 2-fluorobenzoic acid and the 5-fluorocytosine chain (right) in the molecular complex of 5-fluorocytosine and 2-fluorobenzoic acid.

The benzoic acid molecule here is not co-planar with the 5-fluorocytosine chain in the same manner as the 5-fluorocytosine and benzoic acid molecular complex. In this case the entire 2-fluorobenzoic acid molecule is twisted out of the plane in which the homo-base paired chain lies (Figure 3.63).

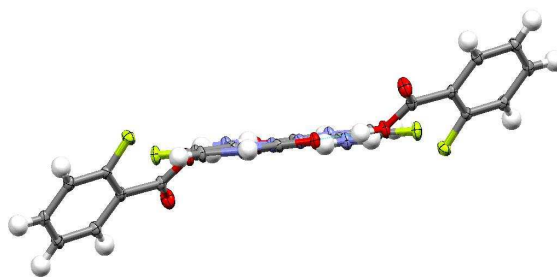


Figure 3.63. Planarity of the 5-fluorocytosine chain and the twisting of the 2-fluorobenzoic acid with respect to this in the molecular complex of 5-fluorocytosine and 2-fluorobenzoic acid.

The presence of a fluorine atom on both molecules in the structure helps to produce a stronger link between the 2-fluorobenzoic acid molecules and the chains to which they are directly hydrogen bonded (Figure 3.64). This increased connectivity comes from the F...F interaction that was not possible in the related 5-fluorocytosine and benzoic acid structure.

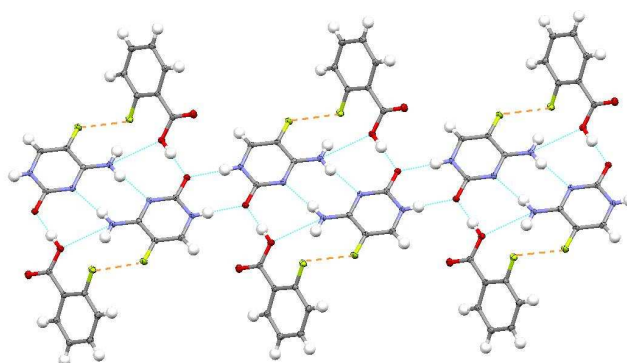


Figure 3.64. Fluorine interactions between the 5-fluorocytosine chain and the tethered 2-fluorobenzoic acid molecules in the molecular complex of 5-fluorocytosine and 2-fluorobenzoic acid.

The fluorine atoms also play a role in linking parallel chains (Figure 3.65). These chains are connected via a fluorine interaction between the fluorine of a 2-fluorobenzoic acid molecule and the fluorine of a 5-fluorocytosine in the chain directly below. This F...F distance is 2.902 (1) Å and is less than the summed van der Waals radii of two fluorine atoms. However, this interaction may just be a consequence of other interactions such as stacking interactions pulling the chains closer together. Further links between the chains are provided in the form of weak C-H...F hydrogen bonds from the hydrogen atom on the ring carbon in the 5 position on the 2-fluorobenzoic acid to the fluorine atom of a 5-fluorocytosine molecule. This has a C...F length of 3.574 (1) Å.

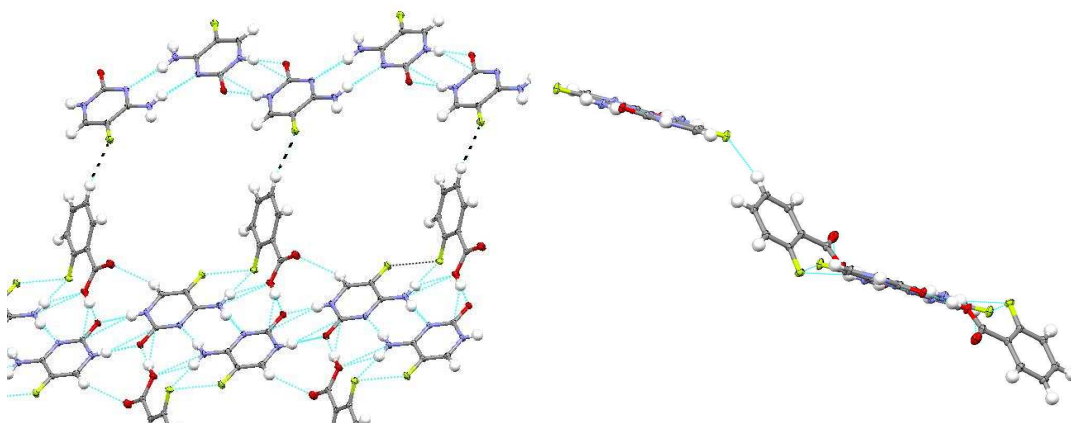


Figure 3.65. Fluorine interactions linking separate chains (left) and the orientation of the chains connected via these fluorine interactions (right) in the molecular complex of 5-fluorocytosine and 2-fluorobenzoic acid.

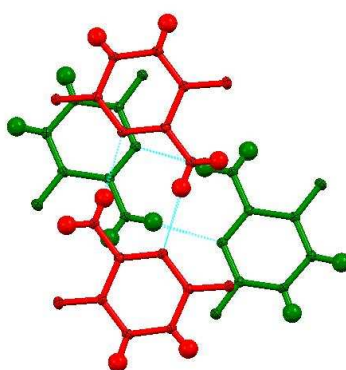


Figure 3.66. Stacking interactions between separate 5-fluorocytosine homo-base paired chains in the molecular complex of 5-fluorocytosine and 2-fluorobenzoic acid.

There are once again base stacking interactions in this structure (Figure 3.66), however in this case there is only one stacking distance. The stacking interactions found here are staggered in nature as the chains do not run in the same direction and do not overlay well. The stacking interactions found were seen to be of approximate length 3.221 Å.

3.6.3. 1:1 Molecular Complex of 5-Fluorocytosine and 3-Fluorobenzoic acid Solvate

5-Fluorocytosine and 3-fluorobenzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.5mm x 0.5mm x 0.4mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 3.14.

The molecular complex forms in a 1:1 ratio of cytosine to benzoic acid. There are two molecules in the asymmetric unit, and the crystal belongs to the space group P-1 with four

molecules in the unit cell. There is no hydrogen transfer in this structure (Figure 3.67). The 3-fluorobenzoic acid molecule exhibits orientational disorder over two positions.

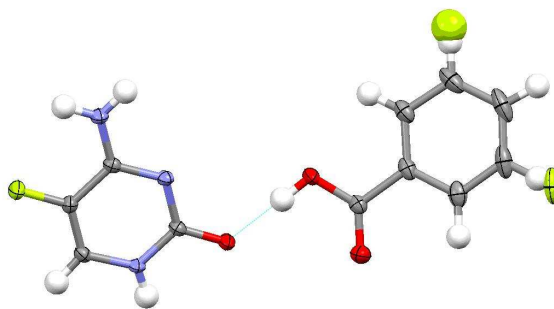


Figure 3.67. The asymmetric unit of the molecular complex of 5-fluorocytosine and 3-fluorobenzoic acid.

As a result of the structure having no hydrogen transfer, the base pairing is again of the pseudo-Hoogsteen type, again forming chains of 5-fluorocytosine with alternating base pairing (Figure 3.68). The N...N distances of the type A base pair are both of equivalent length of 2.994(2) Å making them of moderate strength. There is an inversion centre in the middle of the base pair so the bonds are of equal length. The other base-pairing motif involves the oxygen on the cytosine and the protonated nitrogen in the ring of the cytosine. The hydrogen bonds are once again of moderate strength and have an N...O distance of 2.812(2) Å and are equal once again as there is an inversion centre in the middle of the base-pairing motif.

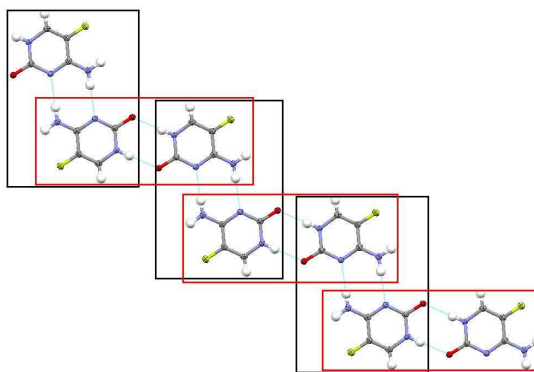


Figure 3.68. The alternating type A (black) and type B (red) base pairing in the molecular complex of 5-fluorocytosine with 3-fluorobenzoic acid.

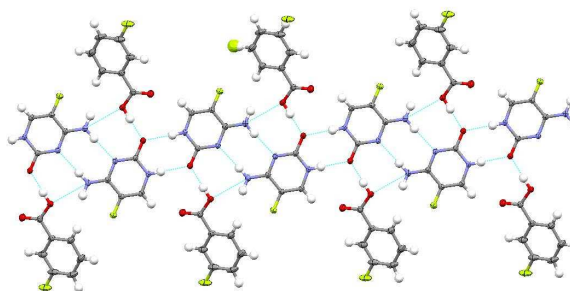


Figure 3.69. Extended homo-base paired chain with 3-fluorobenzoic acid molecules connected to the 5-fluorocytosine chain in the molecular complex of 5-fluorocytosine and 3-fluorobenzoic acid.

The 3-fluorobenzoic acid molecule is connected to the 5-fluorocytosine chain in a similar manner to that of the benzoic acid and 2-fluorobenzoic acid complexes (Figure 3.69). There is an O-H...O hydrogen bond from the hydroxyl oxygen of the carboxyl group to the oxygen on the 5-fluorocytosine. This hydrogen bond is of moderate strength and has an O...O distance of 2.594(2) Å, a similar length to that observed in the 2-fluorobenzoic acid molecular complex. In addition, there is an N-H...O hydrogen bond to this hydroxyl oxygen from the amine group of the 5-fluorocytosine. This hydrogen bond is of moderate strength and has an N...O length of 2.898(2) Å. There is also a weak C-H...F hydrogen bond between an aromatic hydrogen atom on the 3-fluorobenzoic acid and the fluorine atom on the 5-fluorocytosine molecule in the layer below of C...F length 3.478(1) Å.

The 3-fluorocytosine molecule is 75:25 disordered over two positions. This was identified by considering the peak height in a Fourier difference map and the bond distance to the second peak corresponding to the minor component. The proportion of disorder was allowed to freely refine but restraints were used to force the C-F distance for the minor component to take a more realistic value. It was possible to refine the thermal parameter for the major component anisotropically, however, the minor component could only be refined isotropically. It should also be noted that the thermal ellipsoids for the carbon atoms in the benzene ring around this disorder appear to be elongated indicating some libration or the potential for the ring to shift between two positions as a consequence of this disorder.

The framework of the structure is completed through C-H...F hydrogen bonds of H...F distance of 3.485(2) Å between separate chains within the same plane involving the 75% occupied F atom on the 3-fluorobenzoic acid molecule (Figure 3.70). This constructs a hydrogen-bonding pattern that resembles a base-pair motif and there is an inversion centre located in the centre of this motif.

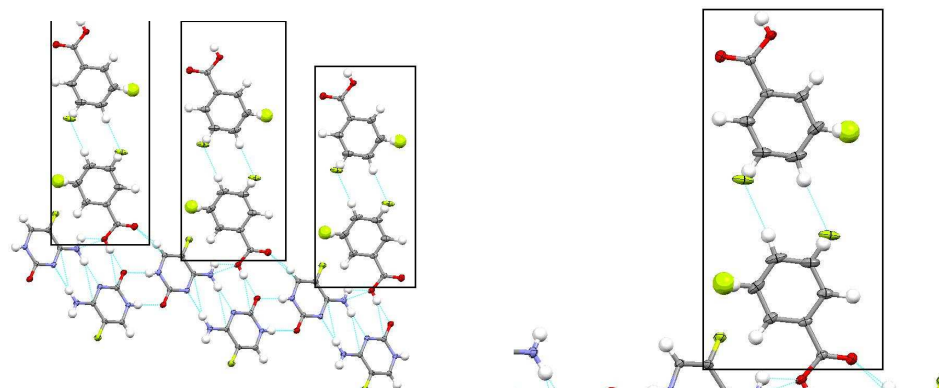


Figure 3.70. Fluorine interactions linking separate chains in the molecular complex of 5-fluorocytosine and 3-fluorobenzoic acid.

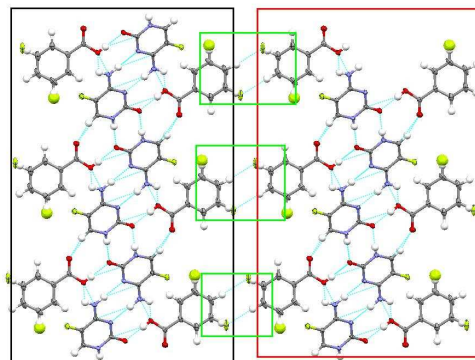


Figure 3.71. Separate chains linked via hydrogen bonds in the molecular complex of 5-fluorocytosine and 3-fluorobenzoic acid (green represents the hydrogen bonds that make use of the fluorine atoms and the red and black represent separate chains).

The structure as a whole is layered and forms a ladder structure with the 3-fluorobenzoic acid molecules making up the rungs. This framework of 5-fluorocytosine base paired chains and hydrogen bonded 3-fluorobenzoic acid molecules results in a large channel running along the *a*-axis. A large amount of residual electron density is located in the channels but it was not possible to resolve the disordered solvent contained within them. The channels could contain either methanol or water and there are no strong hydrogen bonding interactions possible between the solvent in the channel and the framework. The program SQUEEZE [132] was therefore used to remove this electron density and estimate the amount of residual electron density contained within these channels. In the case of the 3-fluorobenzoic acid and 5-fluorocytosine molecular complex there were 36 unaccounted electrons present in the void in the complex. The crystal structure was grown in methanol, therefore it is likely that methanol and/or water are contained in the channel. The 36 electrons could correspond to 2 molecules of methanol per void or 3 molecules of water or a combination of the two.

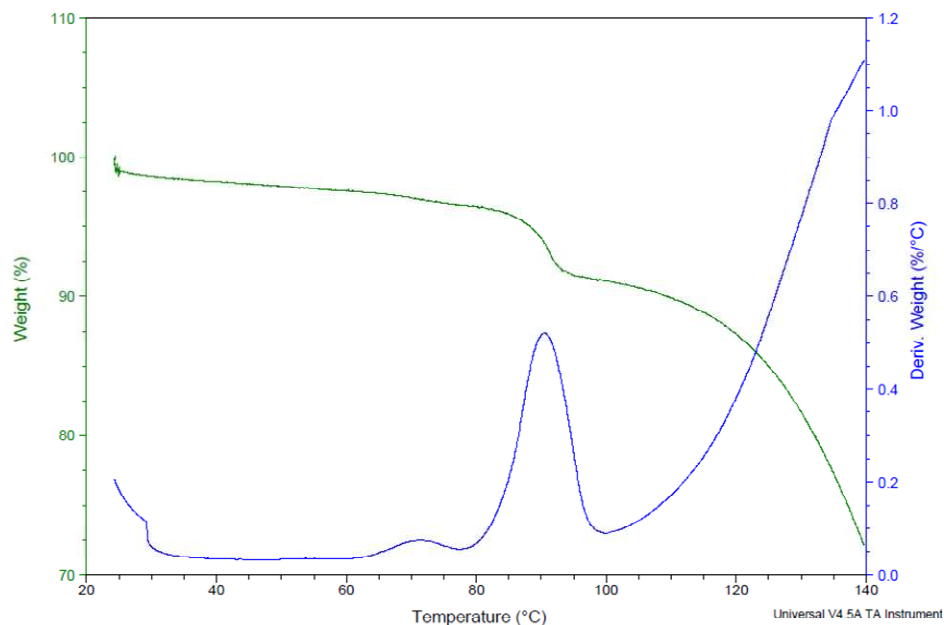


Figure 3.72. TGA thermogram of the molecular complex of 5-fluorocytosine and 3-fluorobenzoic acid.

The properties of the solvent contained within these channels was investigated by DSC and TGA. The TGA thermogram clearly shows a step in the weight loss between 80-100°C and a much smaller dip at around 70°C (Figure 3.72). This would suggest that at least some of the solvent is removed from the sample prior to full decomposition. The DSC thermogram also supports the hypothesis that the disordered solvent is lost before the sample melts as there is a small peak at approximately 100°C and then a larger peak at around 150°C (Figure 3.73). This might suggest that the solvent is released first at 102.5°C and then the sample melts at the second peak at 148.4°C.

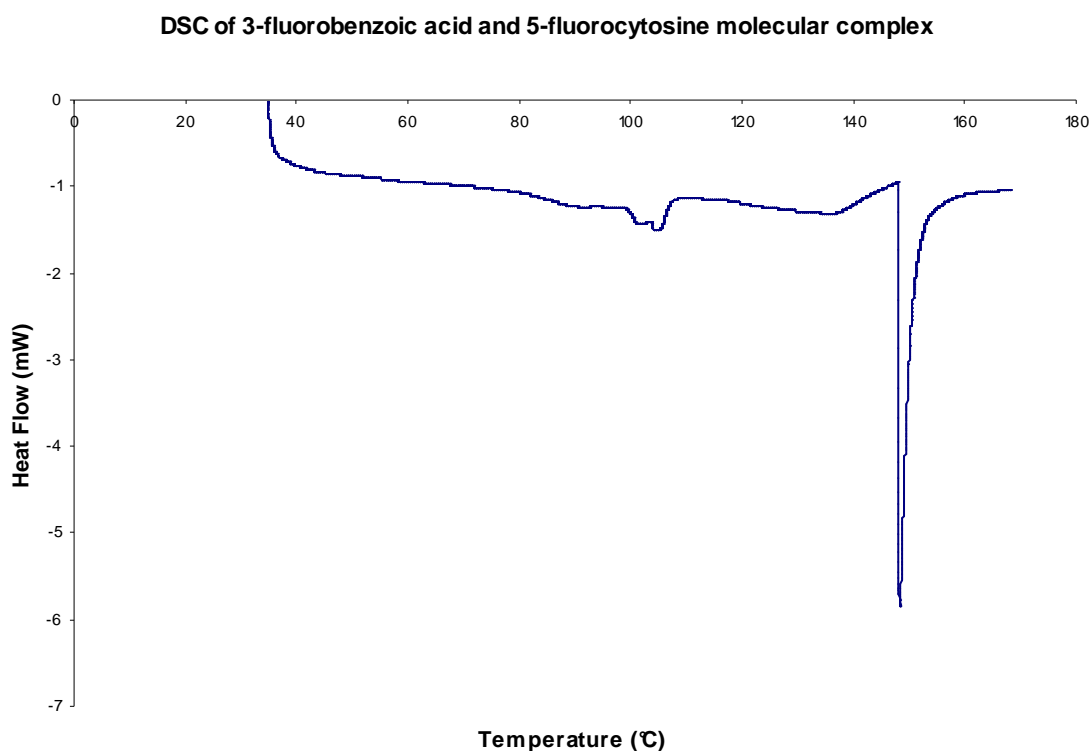


Figure 3.73. DSC thermogram of the molecular complex of 5-fluorocytosine and 3-fluorobenzoic acid.

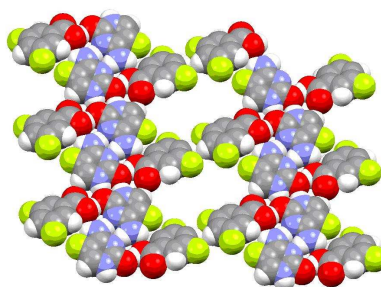


Figure 3.74 Spacefill representation of the molecular complex of 5-fluorocytosine and 3-fluorobenzoic acid showing the presence of a significant void at repeating intervals throughout the structure.

The channel is approximately 7 Å at its smallest diameter and approximately 12 Å at its largest diameter (Figure 3.74). The disordered fluorine is also significant when considering the pores as the 25% occupied F atom points into the void and would be available to interact with the solvent contained there. All interactions with the solvent would, however be weak in nature. If the F atom on both of the neighbouring 3-fluorobenzoic acid molecules constructing the rungs of the ladder were pointing in to the void, there would be no strong link between the 5-fluorocytosine chains. It is therefore likely that when the fluorine atom on one molecule is pointing into the centre of the void, the other fluorine on the neighbouring molecule must be making the link between the separate sides in the form of the C-H...F hydrogen bond. This would mean that the “base-paired” motif between the

separate chains cannot exist 25% of the time as only one weak hydrogen bond will be formed between the two chains.

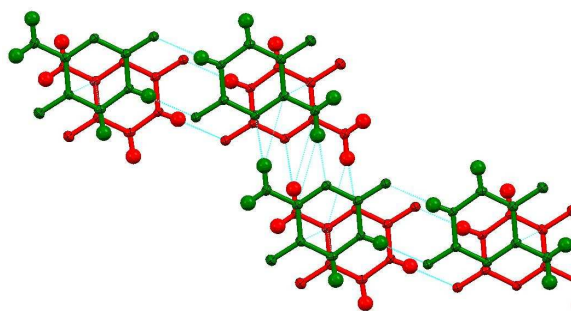


Figure 3.75. Stacking interactions between 5-fluorocytosine chains in the molecular complex of 5-fluorocytosine and 3-fluorobenzoic acid.

The structure once again contains stacking interactions with the cytosine chains packing on top of each other. The relative positions of the cytosine chains show a stagger of approximately half a cytosine molecule (Figure 3.75). The stacking interactions can be classified as base-stacking interactions and the distance between the layers is approximately 3.278 Å.

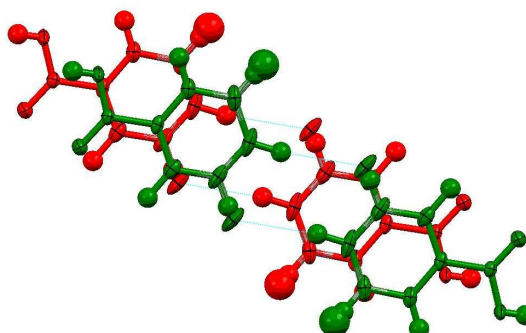


Figure 3.76. Stacking interactions between 3-fluorobenzoic acid molecules in the molecular complex of 5-fluorocytosine and 3-fluorobenzoic acid.

There are also stacking interactions between 3-fluorobenzoic acid molecules which are again staggered with respect to one another with an approximate spacing of 3.389 Å between the layers.

3.6.4. 1:1 Molecular Complex of 5-Fluorocytosine and 4-Fluorobenzoic acid Solvate

5-Fluorocytosine and 4-fluorobenzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.5mm x 0.5mm x 0.4mm. Data were collected on a

Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 3.14.

The molecular complex forms in a 1:1 ratio of 5-fluorocytosine to 4-fluorobenzoic acid. There are two molecules in the asymmetric unit, and the crystal belongs to the space group P-1 with four molecules in the unit cell. There is no hydrogen transfer (Figure 3.77). The structure is isostructural to that of the 3-fluorobenzoic acid 5-fluorocytosine molecular complex.

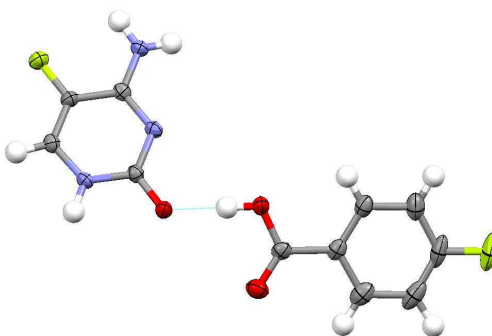


Figure 3.77. Asymmetric unit of the molecular complex of 5-fluorocytosine and 4-fluorobenzoic acid.

The two different pseudo-Hoogsteen base-pair motifs are again found with N-H...N distances of 2.999(3) Å and N-H...O distances of 2.804(4) Å with an inversion centre in the middle of each motif (Figure 3.78). These generate chains of 5-fluorocytosine molecules (Figure 3.79).

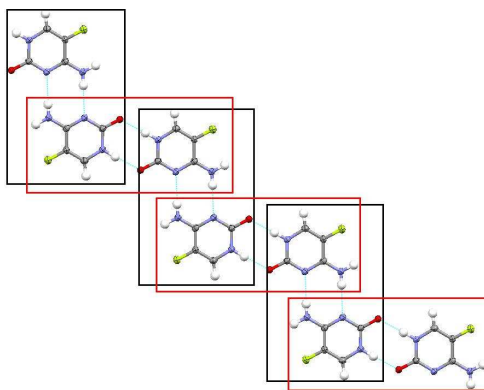


Figure 3.78. The alternating type A and type B base pairing in the molecular complex of 5-fluorocytosine with 4-fluorobenzoic acid.

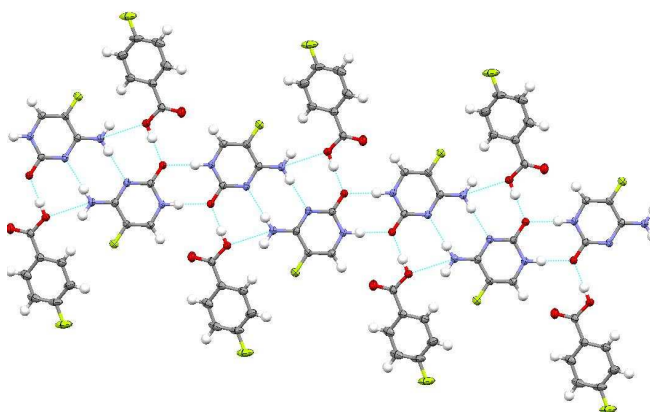


Figure 3.79. Extended homo-base paired chain of 5-fluorocytosine with 4-fluorobenzoic acid molecules connected to the chain in the molecular complex of 5-fluorocytosine and 4-fluorobenzoic acid.

The 4-fluorobenzoic acid molecular complex behaves in a similar manner to the 3-fluorobenzoic acid with respect to the cytosine chain. There is an O-H...O hydrogen bond from the hydroxyl oxygen of the carboxyl group to the oxygen on the 5-fluorocytosine. This bond is of moderate strength and is of length 2.597(2) Å, similar to that observed in the 2-fluorobenzoic acid molecular complex. In addition, there is an N-H...O hydrogen bond also to this hydroxyl oxygen from the amine group of the cytosine. This bond is of moderate strength and is of length 2.901(2) Å. There is also a C-H...F hydrogen bond between an aromatic hydrogen atom on the benzoic acid and the fluorine atom on the 5-fluorocytosine on the layer below which is of weak strength and has a C...F length of 3.500(2) Å. The framework is completed by C-H...F hydrogen bonds connecting 4-fluorobenzoic acid molecules in separate, parallel chains (Figure 3.80). Just as in the 3-fluorobenzoic acid molecular complex, this constructs a hydrogen bonding pattern that resembles a base-pair motif. The C...F distance of these weak hydrogen bonds is 3.490(2) Å with the H...F distance being 2.545(1) Å. This again leads to layers and a framework that resembles a ladder with the benzoic acids making up the rungs (Figure 3.81).

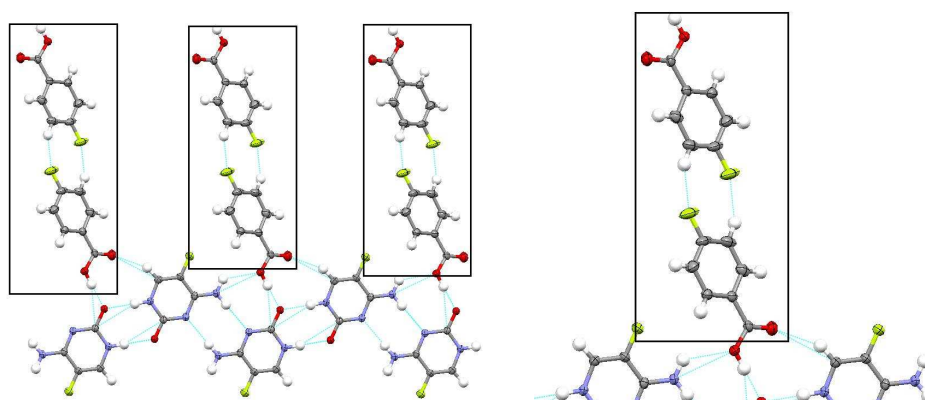


Figure 3.80. Fluorine interactions linking separate chains in the molecular complex of 5-fluorocytosine and 4-fluorobenzoic acid.

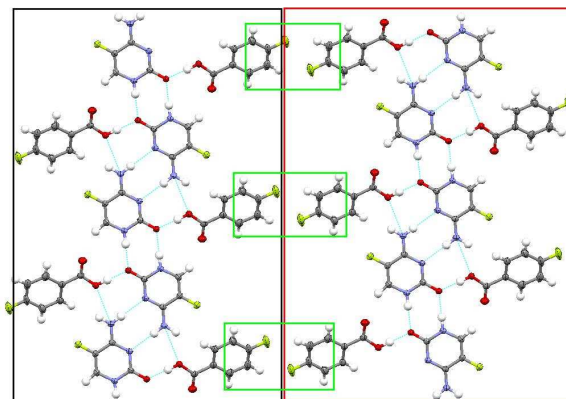


Figure 3.81. Separate chains linked via weak hydrogen bonds in the molecular complex of 5-fluorocytosine and 4-fluorobenzoic acid (green represents the hydrogen bonds that make use of the fluorine atoms and the red and black represent separate chains).

A channel penetrates through the structure along the *a*-axis in the same manner as that observed for the 3-fluorobenzoic acid 5-fluorocytosine molecular complex. A large amount of residual electron density was located in the channel but it was not possible to model the disordered solvent. Again, the likely solvent molecules are methanol or water. SQUEEZE [132] was therefore used to estimate the amount of residual electron density and to remove this electron density for the final refinement. There were 32 unaccounted for electrons present in the void in the structure. This again corresponds to approximately two methanol molecules or three water molecules per void.

DSC and TGA were used to investigate the nature of the solvent contained within the channels and to investigate the structure itself. The TGA (Figure 3.82) and DSC (Figure 3.83) results are less clear with regards to the 4-fluorobenzoic acid molecular complex when compared to those obtained for the 3-fluorobenzoic acid structure. In this structure there is a shoulder in the TGA and a peak in the DSC at roughly 100°C. This may suggest that the solvent is once again being lost; the TGA shows a continuous loss of mass from the start of the heating procedure, with a much less defined shoulder. The DSC thermogram is also inconclusive with several events apparent with a sharp peak at 101.5°C and then a very broad event up until 180°C. Further work would be required on this system before any conclusions can be drawn.

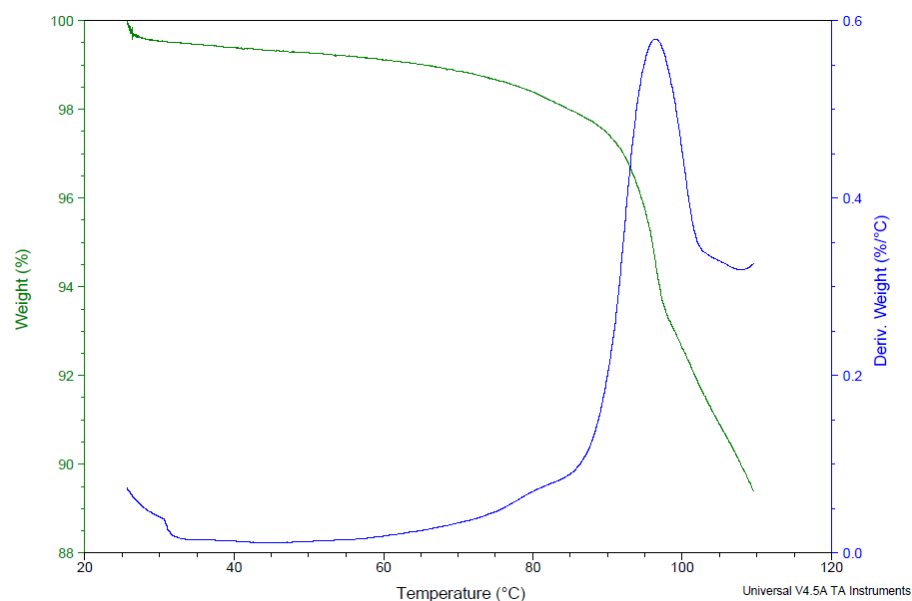


Figure 3.82. TGA thermogram of the molecular complex of 5-fluorocytosine and 4-fluorobenzoic acid

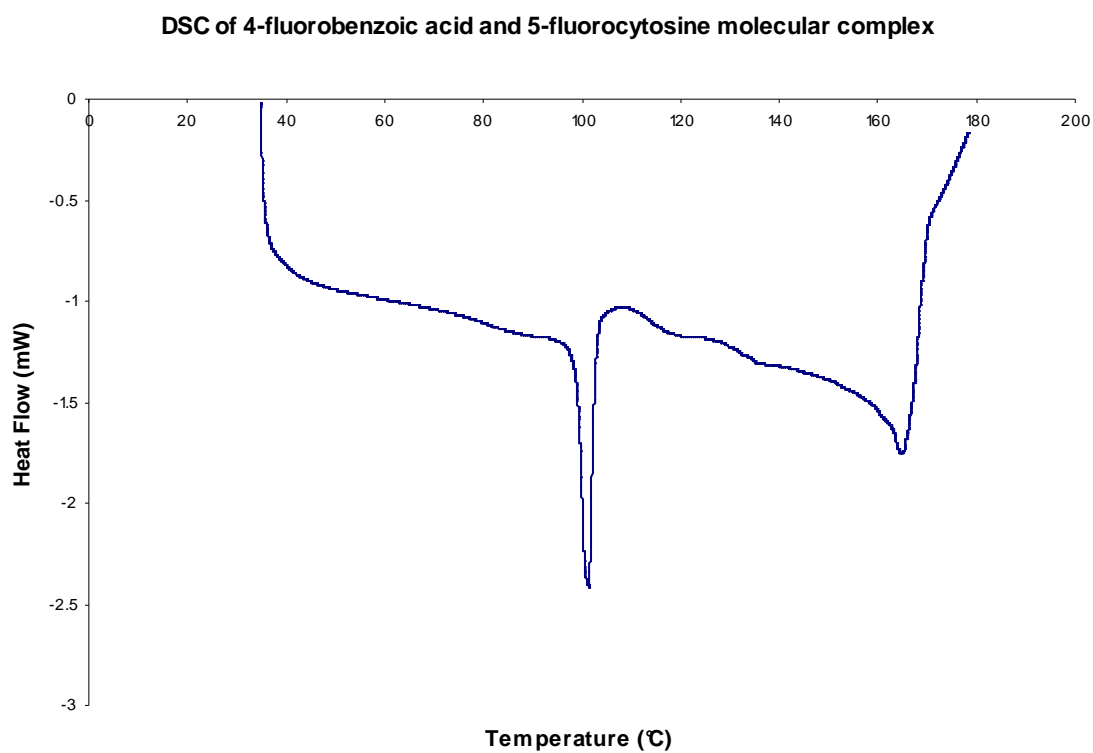


Figure 3.83. DSC thermogram of the molecular complex of 5-fluorocytosine and 4-fluorobenzoic acid.

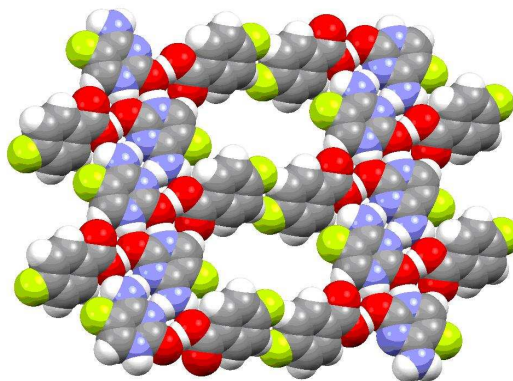


Figure 3.84. Spacefill representation of the molecular complex of 5-fluorocytosine and 4-fluorobenzoic acid showing the significant void.

The channel is a significant proportion of the structure as a whole (Figure 3.84) with a minimum diameter of approximately 9 Å and a maximum diameter of approximately 13 Å at its largest distance.

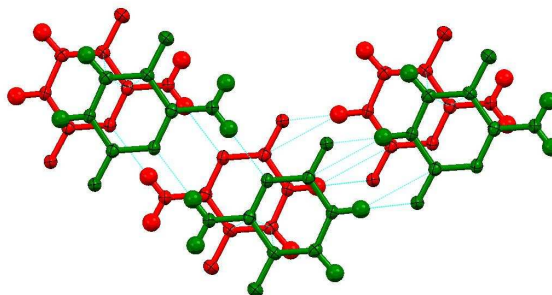


Figure 3.85. Stacking interactions found between 5-fluorocytosine chains in the molecular complex of 5-fluorocytosine and 4-fluorobenzoic acid.

In this structure, in the same manner as the previous structure, there are base-stacking interactions present between two separate chains of homo-base paired cytosine chains. These chains are once again staggered with respect to one another (Figure 3.85). The stacking interactions here are base-stacking interactions and the distance between the layers is approximately 3.280 Å.

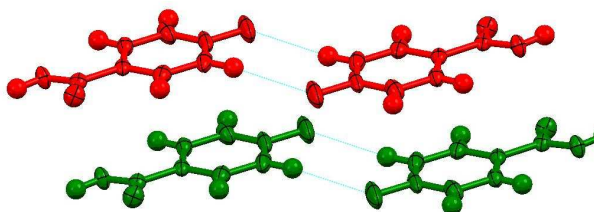


Figure 3.86. Stacking interactions found between 4-fluorobenzoic acid molecules in the molecular complex of 5-fluorocytosine and 4-fluorobenzoic acid.

Similarly to the 3-fluorobenzoic acid structure there are stacking interactions between 4-fluorobenzoic acid molecules (Figure 3.86). The molecules are staggered with respect to each other and the spacing between layers is approximately 3.412 Å.

Table 3.14 Crystal structure information for the molecular complexes of 5-fluorocytosine with benzoic acid and 2-,3- and 4-fluorobenzoic acid

Compound	5-Fluorocytosine and benzoic acid	5-Fluorocytosine and 2-fluorobenzoic acid	5-Fluorocytosine and 3-fluorobenzoic acid solvate	5-Fluorocytosine and 4-fluorobenzoic acid solvate
Formula	C11 H10 N3 O3 F1	C11 H9 N3 O3 F2	C11 H9 N3 O3 F2	C11 H9 N3 O3 F2
Crystallisation Conditions	Methanol, Room Temperature	Methanol, Room Temperature	Methanol, Room Temperature	Methanol, Room Temperature
Molecular weight / g mol⁻¹	251.21	269.20	269.20	269.20
Temperature (K)	100 K	100 K	100 K	100 K
Space Group	P2 ₁ /n	C2/c	P-1	P-1
<i>a</i> (Å)	9.0490(9)	28.7401(17)	3.6572(1)	3.6565(2)
<i>b</i> (Å)	5.3698(5)	10.7454(6)	9.5056(2)	9.5304(5)
<i>c</i> (Å)	22.4839(21)	7.1689(4)	19.3436(4)	19.3970(9)
<i>α</i> (°)	90	90	102.392(1)	77.867(2)
<i>β</i> (°)	93.031(6)	94.910(3)	90.146(1)	86.013(2)
<i>γ</i> (°)	90	90	97.157(1)	83.576(3)
Volume (Å³)	1090.99(4)	2205.80(6)	651.38(2)	655.96(4)
<i>Z</i>	4	8	2	2
<i>θ</i> range (°)	1.8-28.3	1.4-34.0	1.1-27.5	2.2-27.5
Completeness	99.80%	96.30%	99.50%	99.60%
Reflections Collected	34073	84038	14883	17805
Independent	2686	4350	2988	3001
Refln (obs.I>2 <i>σ</i> (I))	2218	3765	2632	1999
<i>R</i>_{int}	0.0447	0.0381	0.0335	0.0630
Parameters	203	208	209	208
Goof on F²	1.016	1.054	1.060	1.059
<i>R</i>₁ (Observed)	0.036	0.037	0.039	0.054
<i>R</i>₁ (all)	0.047	0.043	0.044	0.095
<i>wR</i>₂ (all)	0.097	0.106	0.118	0.128

3.7 Comparison of the Molecular Complexes of 5-Fluorocytosine with Benzoic Acids and their Fluorinated Derivatives

3.7.1. Structural Comparison of Molecular Complexes Formed Between 5-Fluorocytosine and Benzoic Acid and its Mono-Substituted Fluoro Derivatives

Some common trends in the series of related 5-fluorocytosine and benzoic acid molecular complexes can be identified, which are consistent in all four structures. Firstly, there is no hydrogen transfer to the unprotonated nitrogen in the 5-fluorocytosine molecule in any of the structures obtained. Consideration of the ΔpK_a values suggests that this is to be expected as all of these are less than zero (Table 3.15). Secondly, all four structures in this series produce a molecular complex with a crystallisation ratio of 1:1. Two out of the four structures do not contain any solvent in the structure, the molecular complexes of 5-fluorocytosine with benzoic acid and 2-fluorobenzoic acid, whilst the other two structures are solvates with solvent containing channels formed. All four complexes show the same chains comprised of only 5-fluorocytosine and involving two different base pairing motifs, a combination of type A and type B hydrogen bonding motifs.

The two base-pairing types are homo base-pairing motifs that are made up of two hydrogen bonds forming hydrogen-bonded rings. Type A is made up of two N-H...N hydrogen bonds and type B is made up of two N-H...O hydrogen bonds. When a structure contains neutral cytosine or 5-fluorocytosine, these are the predominant motifs between equivalent parts of two cytosine molecules.

Table 3.15. ΔpK_a values, level of hydrogen transfer and the crystallisation ratio for the molecular complexes of 5-fluorocytosine with benzoic acid and 2-, 3- and 4-fluorobenzoic acid.

Molecular complex	ΔpK_a	Hydrogen Transfer with respect to cytosine molecules	Ratio
5-fluorocytosine and benzoic acid	-0.944	None	1:1
5-fluorocytosine and 2-fluorobenzoic acid	-0.010	None	1:1
5-fluorocytosine and 3-fluorobenzoic acid	-0.600	None	1:1
5-fluorocytosine and 4-fluorobenzoic acid	-0.890	None	1:1

3.7.2. Common Hydrogen Bonding Motifs

With all the structures exhibiting no hydrogen transfer then the options available for the 5-fluorocytosine molecule to base-pair in the four separate structures are the same. All four molecular complexes show the same formation of cytosine chains with alternating type A and B hydrogen bonding motifs (Figure 3.87). The presence of fluorine does not affect these primary hydrogen bonding motifs. The differences in the structure are more to do with the orientation of the benzoic acid molecules and the subtle interactions that take place between separate chains in the structures.

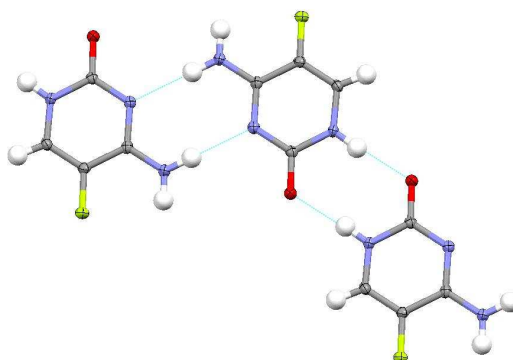


Figure 3.87. Type A and type B hydrogen bonding motifs seen in the extended homo-base paired chains in the molecular complexes of 5-fluorocytosine with benzoic acid and 2-,3- and 4-fluorobenzoic acid.

3.7.3. The role of the benzoic acid molecule and fluorine interactions

In each of the four molecular complexes, benzoic acid molecules are tethered to the cytosine chain through a single hydrogen bond between the carboxylic acid group of the benzoic acid and the carbonyl group of the 5-fluorocytosine molecule. There are two significant effects of fluorine substitution on the benzoic acid molecule: firstly, the orientation of the benzoic acid molecule relative to the cytosine chain; and secondly the torsion angle of the carboxylic acid group relative to the benzene ring of the benzoic acid molecule itself (Table 3.16).

Table 3.16. Angle of the benzoic acid molecule with respect to the 5-fluorocytosine homo-base paired chains for the molecular complexes of 5-fluorocytosine with benzoic acid and 2-,3- and 4-fluorobenzoic acid

	Benzoic acid	2-Fluorobenzoic acid	3-Fluorobenzoic acid	4-Fluorobenzoic acid
Angle between ring and chain	24.43	50.09	28.99	29.87

In general, there is little variation in the angle of the benzene ring of the benzoic acid relative to the 5-fluorocytosine chain. The exception is the 2-fluorobenzoic acid molecular complex where there is a significantly larger angle. This is the only structure that has fluorine-fluorine interactions between the 5-fluorocytosine molecules and the benzoic acid molecules (Figure 3.88, Figure 3.89), and this is enabled by the inherently close positioning of the two, a consequence of the 2-substitution on the benzoic acid. The distance of this fluorine fluorine interaction is 2.721(2) Å, which is significantly shorter than the sum of the van der Waals radii of 2.94 Å. There is also an additional inter-chain N–H...F weak hydrogen bond with the same fluorine which is of N...F distance of 3.1492(9) Å. There is also a C–H...F weak hydrogen bond from the hydrogen in the 5 position of the benzoic acid to the fluorine on the cytosine of a neighbouring chain and this bond is of C...F 3.574(1) Å. The combination of these three fluorine interactions can be considered to be the reason behind the benzoic acid ring orientation being at a larger angle to the cytosine chain than in the other structures.

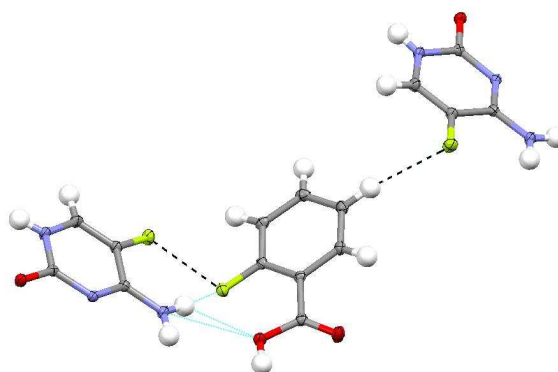


Figure 3.88. Fluorine interactions linking separate chains in the molecular complex of 5-fluorocytosine and 2-fluorobenzoic acid.

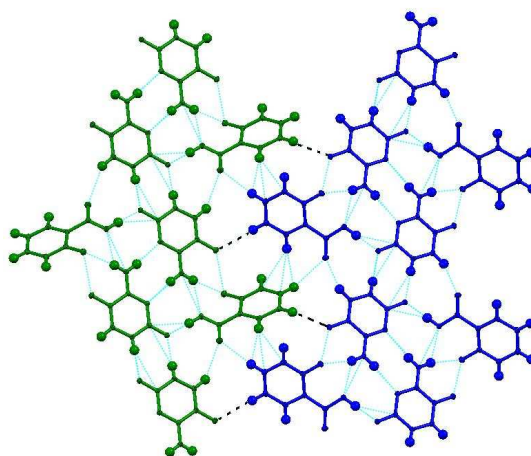


Figure 3.89. Fluorine interactions linking separate chains in the molecular complex of 5-fluorocytosine and 2-fluorobenzoic acid showing extended chains (one represented by blue and one by green).

There are also fluorine-fluorine interactions between layers (Figure 3.90). These distances are 2.902(2) Å and are shorter than the sum of the van der Waals radii (2.94 Å), although much weaker than the other fluorine-fluorine interactions.

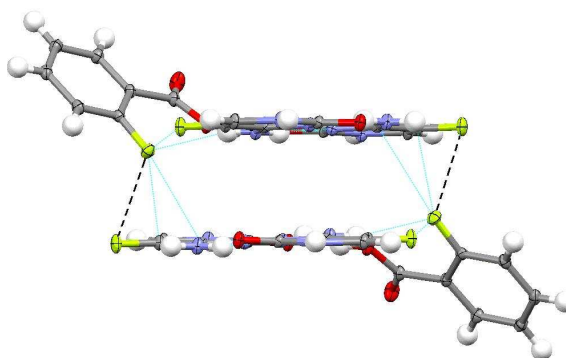


Figure 3.90. Fluorine interactions between the layers in the molecular complex of 5-fluorocytosine and 2-fluorobenzoic acid.

There is also $\pi \cdots \pi$ stacking between layers. In all of the structures there are base stacking interactions. All of the structures show a staggering of the layers which are base-stacked relative to each other with very similar distances between the layers (Table 3.17).

Table 3.17. Comparison of the base stacking interactions for the molecular complexes of 5-fluorocytosine with benzoic acid and 2-,3- and 4-fluorobenzoic acid.

Structure	Base-stacking interactions distances
5-fluorocytosine and benzoic acid	3.268 Å
5-fluorocytosine and 2-fluorobenzoic acid	3.221 Å
5-fluorocytosine and 3-fluorobenzoic acid	3.278 Å
5-fluorocytosine and 4-fluorobenzoic acid	3.280 Å

The other three structures do not have strong fluorine-fluorine interactions between the cytosine chains and the benzoic acid molecules hydrogen bonded to these chains. However fluorine interactions are still crucial to the secondary structural aspects. The 3-fluorobenzoic acid and 4-fluorobenzoic acid molecular complexes are isostructural to one another, perhaps surprisingly. Despite the different position of the fluorine on the benzoic acid ring, the 5-fluorocytosine chains are connected to one another through the same secondary building unit creating channels in which disordered solvent is present.

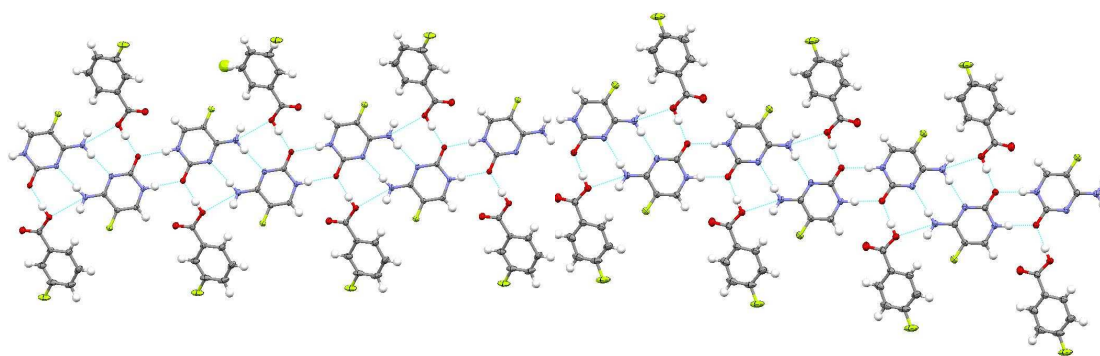


Figure 3.91. Extended chains in the molecular complexes of 3-fluorobenzoic acid and 5-fluorocytosine (left) and 4-fluorobenzoic acid and 5-fluorocytosine (right).

3.7.4. Isostructural molecular complexes of 3-fluorobenzoic and 4-fluorobenzoic acid with 5-fluorocytosine

In the case of these two molecular complexes, the fluorine on the benzoic acid molecule is less sterically constrained and as a result is able to interact with benzoic acid molecules which are tethered to neighbouring chains. A pseudo-base-pair motif is formed between two benzoic acid molecules through two C–H...F hydrogen bonds with an inversion centre located in the centre of the dimeric unit.

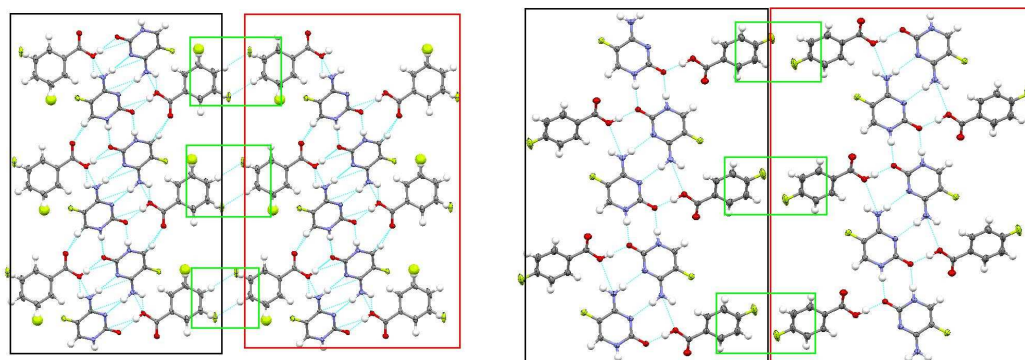


Figure 3.92. Extended chains linked together via fluorine included base-paired motif present in the molecular complexes of 3-fluorobenzoic acid and 5-fluorocytosine (left) and 4-fluorobenzoic acid and 5-fluorocytosine (right).

These pseudo-base-paired two molecule units form the rungs of a ladder connecting two cytosine hydrogen bonded chains (Figure 3.92). Weak hydrogen bonds hold the links together with C...F distances of 3.485(2) in the 3-fluorobenzoic acid molecular complex and 3.490(2) Å in the 4-fluorobenzoic acid molecular. These interactions play a significant role in linking the chains (Figure 3.93, Figure 3.94) and create a porous structure with channels that run along the *a*-axis that is discussed further below.

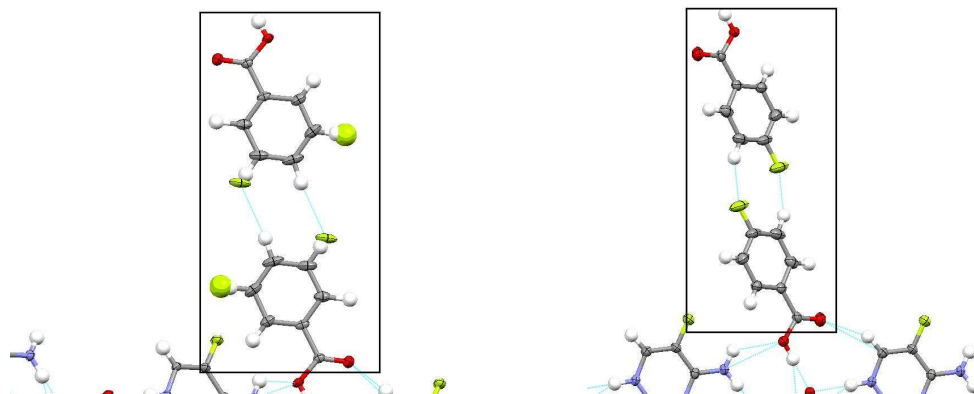


Figure 3.93. Close up of fluorine base paired style motif linking separate chains in the 3-fluorobenzoic acid and 5-fluorocytosine (left) and the 4-fluorobenzoic acid and 5-fluorocytosine molecular complexes.

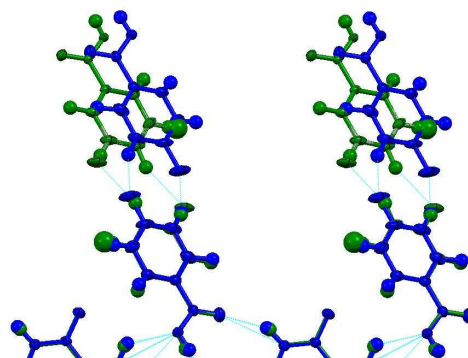


Figure 3.94. Structure overlay of the 3-fluorobenzoic acid and 5-fluorocytosine molecular complex (green) and 4-fluorobenzoic acid and 5-fluorocytosine molecular complex (blue) showing the links between the chains.

The similarity between the two 5-fluorocytosine chains with tethered hydrogen bonded benzoic acid molecules can be clearly seen by viewing a structural overlay of the two (Figure 3.95). However, the change in position of the fluorine from the 3 to the 4 position results in differences between the two structures over longer length scales.

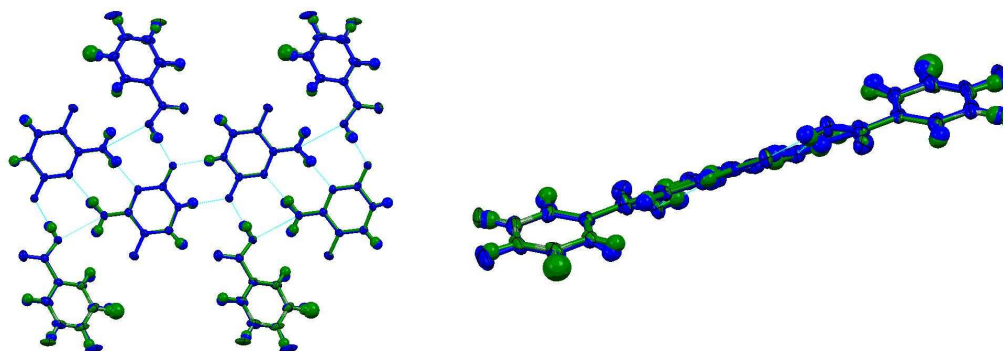


Figure 3.95. Structure overlay of the 3-fluorobenzoic acid and 5-fluorocytosine molecular complex (green) and 4-fluorobenzoic acid and 5-fluorocytosine molecular complex (blue).

In the 4-fluorobenzoic acid molecular complex, the two C–H...F hydrogen bonded benzoic acid molecules are staggered relative to one another whilst in the 3-fluorobenzoic acid

molecular complex this shift is much less pronounced. This is the reason behind the small differences in the two structures. Both structures are porous and contain channels that have disordered solvent within them (Figure 3.96). This disordered solvent could not be modelled and so the program SQUEEZE [132] was used to estimate and remove the electron density associated with the disordered solvent. This is not surprising given the framework of the channels contains only weak hydrogen bond donors and acceptors in the form of aromatic hydrogen and fluorine atoms; any solvent might thus be expected to only be weakly bound. In the case of the 3-fluorobenzoic acid there are approximately 36 residual electrons contained within the pore. The crystals were grown in methanol and thus the pore could contain methanol or water; this equates to either 4 water molecules or 2 methanol molecules. In the case of 4-fluorobenzoic acid there are approximately 32 residual electrons contained within the pore; this could equate to approximately 3 water molecules present or approximately 1.5 methanol molecules per channel. There is also the possibility that there is a mixture of water and ethanol within the channels. The channels in the structure are significant (Table 3.18); the 4-fluorobenzoic acid molecular complex contains a larger pore than that in the 3-fluorobenzoic acid molecular complex.

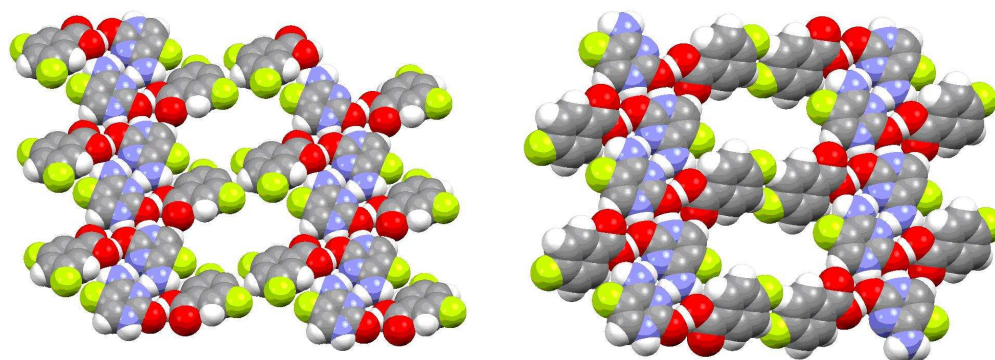


Figure 3.96. Spacefill representation to show the voids present in the molecular complexes of 3-fluorobenzoic acid and 5-fluorocytosine (left) and 4-fluorobenzoic acid and 5-fluorocytosine (right).

Table 3.18. % of the unit cell made up of the void in the molecular complexes of 5-fluorocytosine with 3-fluorobenzoic acid and 4-fluorobenzoic acid.

	3-Fluorobenzoic acid and 5-fluorocytosine	4-Fluorobenzoic acid and 5-fluorocytosine
% of the Unit Cell made up by void	15.4	18.6

Thermogravimetric analysis (TGA) was used to investigate the nature of the solvent contained within the channels. The loss of mass between 80-100°C in the 3-fluorobenzoic acid 5-fluorocytosine molecular complex shows the loss of solvent from the channels prior to decomposition of the framework structure. It could be that the small peak present around the 65-70°C mark could be from methanol leaving the structure and the peak around 80-

100⁰C mark results from water leaving the structure. This result could point to the fact that there would be methanol and water present in the void. The 4-fluorobenzoic acid 5-fluorocytosine molecular complex shows a weight change at around 100⁰C, however, it is unsure if this peak is due to the disordered solvent or merely the start of the decomposition of the sample. The DSC of each molecular complex might point to the fact that the crystal could have the solvent removed and then recover the crystal and retain the porous structure with the solvent no longer present in the voids. However, the results of the 3-fluorobenzoic acid structure are far more comprehensive than the results obtained from the 4-fluorobenzoic acid structure.

3.7.5. Molecular Complex of Benzoic Acid and 5-Fluorocytosine

The benzoic acid 5-fluorocytosine molecular complex is quite different to all of the other structures; the chains are linked together via C–H...F weak hydrogen bonds (Figure 3.97, left) but these interactions are not from a chain lying in the same plane but with a neighbouring chain oriented approximately 90° to the original chain (Figure 3.97, right).

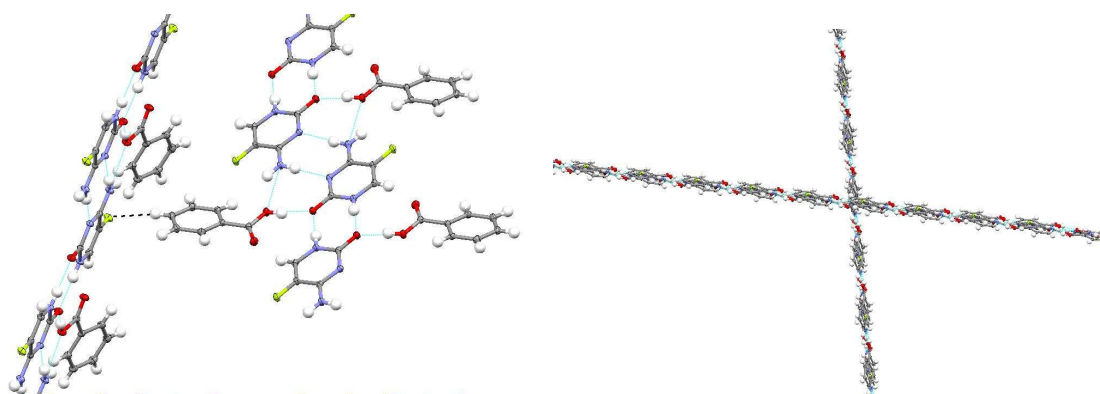


Figure 3.97. Fluorine interactions helping to link separate chains together in the 5-fluorocytosine and benzoic acid molecular complex.

The C...F distance of the interaction is 3.321(2) Å and this can be classed as a weak hydrogen bond. This structure is much more closely packed than the 3-fluorobenzoic acid and the 4-fluorobenzoic acid molecular complexes, with no porous channels formed. Figure 3.98 illustrates the layered nature of the 5-fluorocytosine benzoic acid molecular complex.

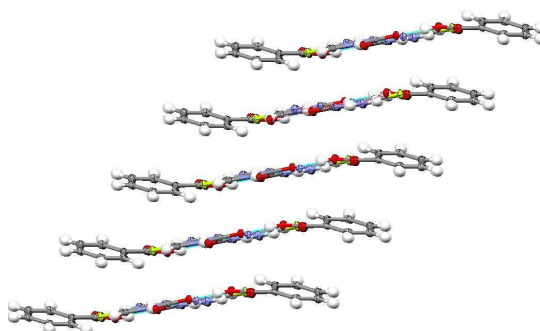


Figure 3.98. Stacked layers in the 5-fluorocytosine and benzoic acid molecular complex.

When comparing the layers in the four structures there appears to be two different classes; one class is the porous structure and the second are closely packed structures. The benzoic acid and 2-fluorobenzoic acid structures are similar to an extent in that both are very close packed structures in comparison to the other structures, and layers can be identified in both; however, the nature of these layers is quite different (Figure 3.99).

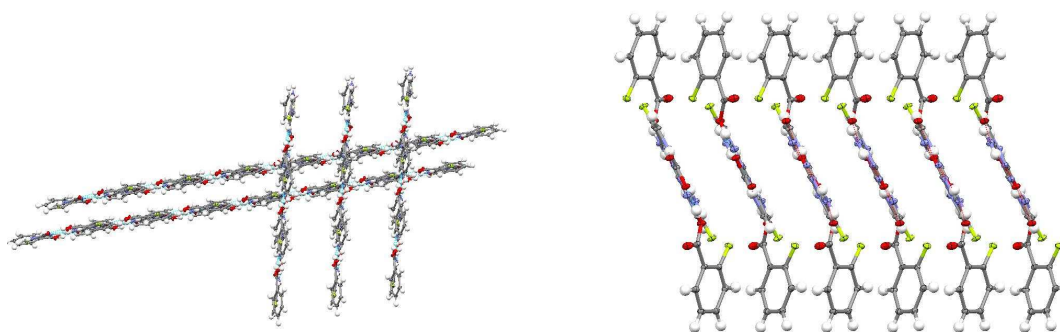


Figure 3.99. Chains running perpendicular to each other in the molecular complex of 5-fluorocytosine and benzoic acid (left) and the stacked layers present in the 5-fluorocytosine and 2-fluorobenzoic acid molecular complex with no void in either structure.

3.3 Conclusions

There are some trends and some preferred bonding patterns that can be identified in the molecular complexes reported in this chapter. One of the most important structure directing trends is to do with hydrogen transfer. In the structures that involve the cytosine molecule it appears that hydrogen transfer readily occurs whilst this is not a common occurrence when dealing with the 5-fluorocytosine molecule. In the molecular complexes involving the cytosine molecule, three out of the four structures have partial transfer (both charged and neutral cytosine molecules present in the structure) and this directs these structures towards forming the pseudo-Watson-Crick hydrogen bonding motif. The remaining structure contains total hydrogen transfer which means the pseudo-Watson-Crick bonding motif is not possible and this limits the structure to forming the hydrogen-bonded ring motif.

The structures containing the 5-fluorocytosine molecule do not contain any hydrogen transfer and as a result are not capable of forming the pseudo-Watson-Crick bonding motif. All four structures here exhibit the same type A and type B hydrogen bonding motifs, which combine alternately to form an extended 5-fluorocytosine chain.

From the work reported in this chapter it becomes apparent that if a fluorine atom is present as a ring substituent, the most common type of interaction in which it will become involved will be weak hydrogen bonds. The fluorine hydrogen bonds appear in several structures and range from acting as a link between layers to being the link between adjacent chains. The fluorine atom can also be involved in several interactions with nearby molecules, with oxygen atoms for example, that add stability to the structure that has been originally constructed via the hydrogen bonded motifs discussed. The fluorine plays an important role in the structures in which it is present but the main interactions will be predominantly dependent on the original bonding motif adopted by the structure.

In this chapter it has been seen that there are stacking interactions present in the structures. The predominant motif of stacking interactions will be staggered with the separate chains not lining up perfectly, which has been seen in the majority of the structures. The actual distances in the base stacking interactions can be seen to be very consistent with the values ranging from 3.1 to 3.3 Å.

Chapter 4. Molecular Complexes of Cytosine and 5-Fluorocytosine with Hydroxybenzoic Acids.

Molecular complexes formed between cytosine and 5-fluorocytosine and the series of mono-substituted hydroxybenzoic acids will be discussed in this chapter. Proton transfer effects and the dominant hydrogen bonding motifs will be described along with an investigation into the effect of fluorination on the molecular complexes obtained.

4.1. 1:1 Molecular Complex of Cytosinium Salicylate

The molecular complex of cytosinium salicylate has previously been reported [133]. The crystal structure will be discussed in detail in the context of the wider series to allow comparison with the related molecular complexes reported in this thesis.

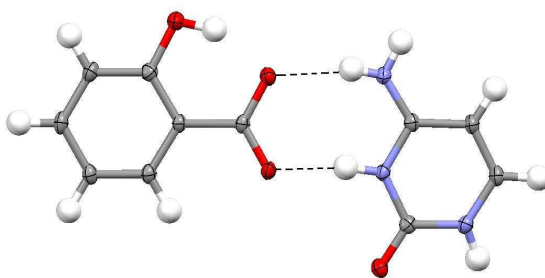


Figure 4.1. The heterodimer formed in the cytosinium salicylate molecular complex.

This molecular complex shows hydrogen transfer from the carboxyl group of the salicylic acid molecule to the unprotonated nitrogen of the cytosine ring. There are therefore two charged species in the structure; the salicylic acid has lost a hydrogen atom to form a salicylate ion and the cytosine has gained a hydrogen atom to create a cytosinium ion. Due to the 100% transfer of the hydrogen, the pseudo Watson Crick hydrogen bond motif cannot form. The next most common hydrogen-bonding pattern found in Chapter 3 in the absence of the pseudo Watson Crick the other possibility for this structure is the type A and C bonding motif between pairs of cytosine/cytosinium molecules and a hydrogen bonded ring between cytosine/cytosinium and co-molecule. The type B hydrogen-bonding motif is also possible and this can occur in combination with the other primary bonding motifs. All of these hydrogen-bonding motifs involve two hydrogen bonds (Section 1.10). In this molecular complex, a hydrogen-bonded ring between the cytosinium and salicylate molecules is formed creating a heterodimer (Figure 4.1), and this hydrogen bonding pattern involves the two areas that are involved in the hydrogen transfer.

The hydrogen bonds in this heterodimer are of N-H...O type and of moderate strength. The first hydrogen bond is between one of the hydrogen atoms on the amine group of the cytosinium and one of the oxygen atoms on the carboxylate group and has an N...O distance of 2.784(2) Å. The second has an N...O distance of 2.724(2) Å, and involves a protonated nitrogen of the cytosinium molecule and the other oxygen in the carboxylate group. There is also an intramolecular hydrogen bond that helps to stabilise the charge on the salicylate molecules between the hydroxyl group and one of the oxygen atoms in the carboxylate group. This hydrogen bond is strong with an O...O distance of 2.542(3) Å.

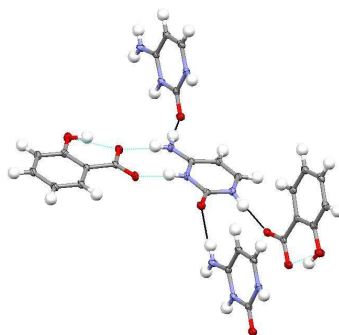


Figure 4.2. The hydrogen bonding environment of the cytosinium ion in the cytosinium salicylate molecular complex.

Each cytosinium molecule is involved in a total of five hydrogen bonds: two make up the heterodimer hydrogen-bonded ring, an additional one is formed to a different salicylate molecule, and two are formed to other cytosinium molecules (Figure 4.2). The two hydrogen bonds between amine groups and carbonyl oxygens of cytosinium molecules are equivalent and of moderate strength with N...O distances of 2.872(1) Å. The cytosinium molecules are oriented approximately perpendicular to one another. These hydrogen bonds create a stepped chain of heterodimers along the *ac* direction creating an L shape (Figure 4.3), which has an angle at the centre of the L of ~90°. These chains are connected to one another through N-H...O hydrogen bonds between a ring nitrogen and the carboxylate group on a neighbouring salicylate molecule (Figure 4.4 (b)). These hydrogen bonds are of moderate strength and of N...O distance of 2.693(2) Å. This creates a herring-bone structure (Figure 4.4).

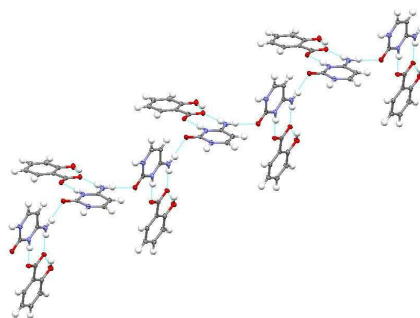


Figure 4.3. Chains of hydrogen bonded heterodimers in the cytosinium salicylate molecular complex.

Figures 4.4, left and 4.5 highlight the presence of base-stacking between these L shaped chains. There is base stacking between the red chain and the blue chain, the green chain and the blue chain and the yellow and green chains, with two different layer spacings (Table 4.1). The salicylate ions are parallel to one another in different planes and the cytosinium molecules are also parallel to one another. However, the hydrogen bonded unit is not planar and so the planes generated by cytosinium molecules are not parallel to those generated by salicylate molecules.

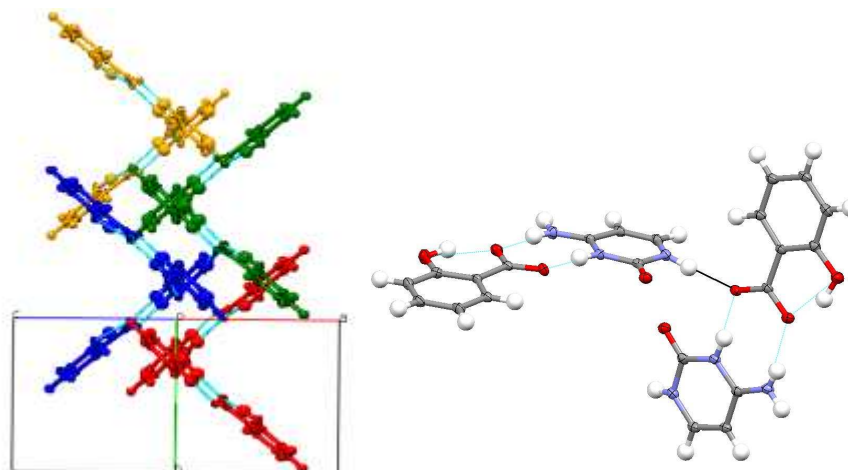


Figure 4.4. Left, the herringbone arrangement of hydrogen bonded L shaped chains. All L shaped units are symmetry related; right, link between separate chains to create the L shape via moderate N-H...O hydrogen bonds.

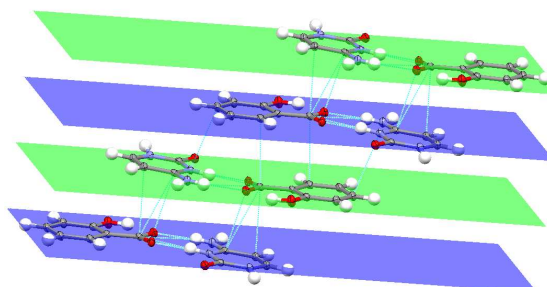


Figure 4.5. Base stacking in the cytosinium salicylate molecular complex. The units are staggered with respect to one another and the planes are generated through salicylate molecules.

Table 4.1. Spacing between layers in the cytosinium salicylate molecular complex

Layers	Distance between the layers (Å)
Blue and Green	3.164
Green and Blue	3.136

The base stacking interactions are short and significant between the layers. The relative positioning of the molecules in the layers of green and yellow compared to green and blue can explain the slight differences between the distances between layers. The base units are staggered relative to one another in separate planes and the salicylate ions stack above the heterodimer hydrogen bonded rings in the shorter spacing and the hydrogen bonded rings stack above one another in the larger spacing with a cytosinium molecule above a salicylate molecule and vice versa in both cases (Figure 4.5). There are only base stacking interactions present between the yellow and green layer and the blue and red layer in the structure due to the orientation of the different layers. Figure 4.6 exemplifies this, with the nearest molecules to the original pair being perpendicular to the original. This means there cannot be any stacking interaction between green and blue. This means that there are only two unique stacking interactions present between the layers and the blue and green layers present in Figure 4.5 show this.

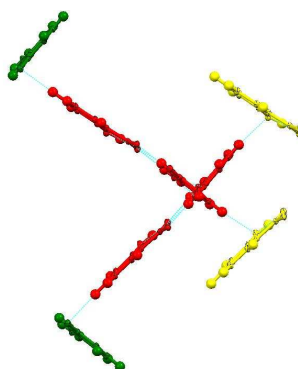


Figure 4.6. The herringbone arrangement of the hydrogen bonded L shaped chains with weak hydrogen bonds linking molecules to either end of the chains.

The L shaped units are also connected to one another by weak C-H...O hydrogen bonds (Figure 4.6). At the tips of the L shaped units, the hydrogen bond is between the hydrogen in the para position to the carboxylate group of the salicylate molecule and the oxygen atom on a nearby cytosinium molecule (yellow dashed lines in Figure 4.6) with a C...O distance of 3.527(3) Å. At the apex of each L shaped unit, there is a second weak hydrogen bond between an aromatic ring carbon of a cytosinium molecule to the oxygen atom of the hydroxy group of a salicylate molecule (blue dashed lines in Figure 4.6). This has a C...O distance of 3.514(3) Å.

Table 4.2. Crystallographic data for the molecular complexes of cytosine and 5-fluorocytosine with salicylic acid and 5-fluorosalicic acid

Compound	Cytosinium salicylate *Data collected by S.M. Harte [133].	5-Fluorocytosinium salicylate	Cytosinium 5- fluorosalicylate	5-Fluorocytosine and 5- fluorosalicic acid hydrate	Cytosine and 5- fluorosalicic acid hydrate
Formula	C11 H11 N3 O4	C11 N3 O4 H10 F1	C11 H10 N3 O4 F1	C15 H17 N6 O7 F3	C16 H15 N6 O7 F1
Crystallisation Conditions	Methanol, Room Temperature	Methanol, Room Temperature	Methanol, Room Temperature	Methanol, Room Temperature	Methanol, Room Temperature
Molecular weight / gmol⁻¹	249.22	267.21	267.2	450.33	422.33
Temperature (K)	100 K	100 K	100 K	100 K	100 K
Space Group	P2 ₁ /n	P2 ₁ /n	P2 ₁ /n	P-1	P-1
a (Å)	9.6425(3)	7.8518(17)	8.3534(7)	8.8996(1)	8.8791(6)
b (Å)	8.4698(3)	10.6457(24)	10.1434(7)	10.3492(2)	10.2957(8)
c (Å)	13.9021(4)	13.3065(29)	13.1203(10)	11.3601(2)	11.1479(10)
α (°)	90	90	90	101.195(1)	98.713(3)
β (°)	109.887(2)	100.041(15)	101.860(4)	108.857(1)	109.424(4)
γ (°)	90	90	90	103.707(1)	104.547(4)
Volume (Å³)	1067.68(6)	1095.23(65)	1087.98(9)	919.30(7)	1079.81(2)
Z	4	4	4	2	2
θ range (°)	2.256-30.034	2.5-33.6	2.6-27.5	2.0-25.0	2.005-30.252
Completeness	99.60%	99.50%	99.60%	99.70%	97.50%
Reflections Collected	11646	34882	19935	16926	10619
Independent	3101	4028	2461	3246	5211
Refln (obs.I>2 σ (I))	2309	2809	1884	2826	3205
R_{int}	0.173	0.109	0.0452	0.0409	0.0452
Parameters	164	212	212	348	212
GooF on F²	0.9502	1.094	1.039	1.034	1.039
R₁ (Observed)	0.065	0.094	0.033	0.031	0.0672
R₁ (all)	0.0829	0.061	0.051	0.039	0.091
wR₂ (all)	0.2015	0.208	0.093	0.083	0.1332

4.2. 1:1 Molecular Complex of 5-Fluorocytosinium Salicylate ^[134]

The molecular complex of 5-fluorocytosinium salicylate has also been previously reported [133,134]. This structure has been redetermined at 100K to allow direct comparison of the hydrogen bonding parameters with the other molecular complexes reported in this thesis and also to confirm the position of the hydrogen atoms at low temperatures.

5-Fluorocytosine and salicylic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.3mm x 0.4mm x 0.4mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. The molecular complex forms in a 1:1 ratio and crystallographic data are given in Table 4.2.

This molecular complex shows 100% hydrogen transfer from the salicylic acid to the 5-fluorocytosine, from the carboxyl group of the salicylic acid to the unprotonated nitrogen of the cytosine ring. There are therefore two charged species in the structure, a 5-fluorocytosinium ion and a salicylate ion. A hydrogen-bonded ring between the 5-fluorocytosine and the salicylic acid is formed generating a heterodimer (Figure 4.7) similar to that found in the cytosinium salicylate molecular complex (Section 4.1).

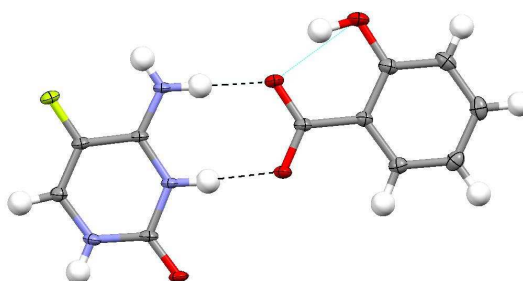


Figure 4.7. 5-Fluorocytosinium salicylate hydrogen bonded heterodimer.

The hydrogen bonds comprising this hydrogen-bonded ring are of N-H...O type and moderate strength. The first hydrogen bond is between one of the hydrogen atoms on the amine group of the 5-fluorocytosinium and one of the oxygen atoms on the carboxylate group and has an N...O distance of 2.751(2) Å. The second hydrogen bond has an N...O distance of 2.770(2) Å, and involves a protonated ring nitrogen of the 5-fluorocytosinium molecule and the other oxygen in the carboxylate group. There is also an intramolecular hydrogen bond that helps to stabilise the charge on the salicylate molecule. This hydrogen bond is strong in strength and is has an O...O distance of 2.550(2) Å.

Similarly to the cytosinium salicylate molecular complex, the 5-fluorocytosinium molecule is involved in a total of five hydrogen bonds: two make up the heterodimer hydrogen bonded ring, an additional one is formed to a different salicylate molecule, and two to two separate 5-fluorocytosinium molecules (Figure 4.8). The salicylate ions are parallel to one another in different planes and the 5-fluorocytosinium molecules are also parallel to one another. However, the hydrogen-bonded unit is not planar and so the planes generated by 5-fluorocytosinium molecules are not parallel to those generated by salicylate molecules.

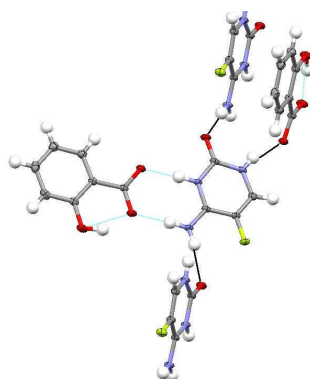


Figure 4.8. The hydrogen bonds involving a 5-fluorocytosinium molecule in the molecular complex of 5-fluorocytosinium salicylate.

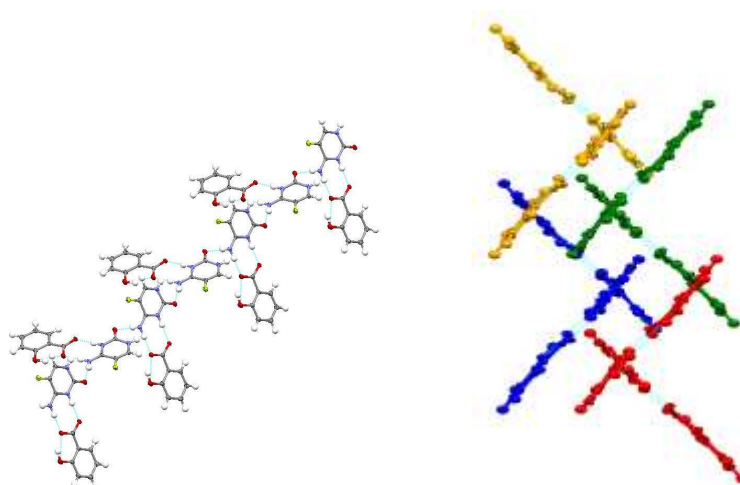


Figure 4.9. Left, chain of hydrogen bonded heterodimers in the 5-fluorocytosinium salicylate molecular complex; right, the herringbone structure of hydrogen bonded L shaped chains.

The hydrogen bonds between the amine groups and the carbonyl oxygens of 5-fluorocytosinium molecules are once again equivalent and of moderate strength with $N\cdots O$ distances of 2.871(1) Å. The 5-fluorocytosinium molecules connected through this hydrogen bond lie approximately perpendicular to one another to create a stepped chain of heterodimers along the *ac* direction again creating an L shape (Figure 4.9), which has an angle of $\sim 117^\circ$ at the centre of the L shape, considerably more obtuse than that found in the cytosinium salicylate molecular complex ($\sim 90^\circ$). These chains are linked to one another via $N-H\cdots O$ hydrogen bonds between a ring nitrogen and the carboxylate group on a neighbouring salicylate molecule (Figure 4.8). These hydrogen bonds are of moderate

strength and have an N...O distance of 2.701(2) Å. This, once again, creates a herring-bone structure (Figure 4.9, right).

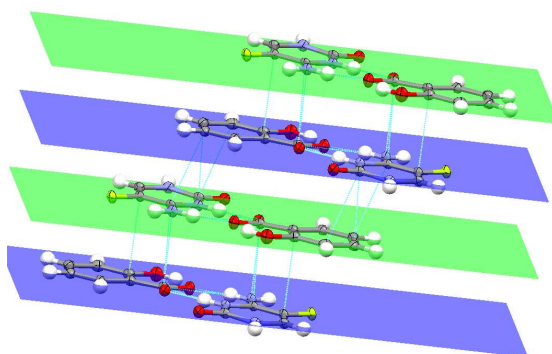


Figure 4.10. Base-stacking in the 5-fluorocytosinium salicylate molecular complex.

Figure 4.10, much like Figure 4.5 for the non-fluorinated structure, helps to illustrate the presence of base-stacking in the structure. The base stacking is clearly evident between the blue and green planes. The separation of these base stacked layers is approximately 3.208 Å and 3.150 Å, representing significant interactions. There is only base stacking interactions present between the yellow and green layer and the blue and red layer in the structure due to the orientation of the different layers. Figure 4.10 exemplifies this, with the nearest molecules to the original pair being perpendicular to the original. This means there cannot be any stacking interaction between green and blue. This means that there are only two unique stacking interactions present between the layers; the blue and green layers present in Figure 4.10 show this. Again the base units are staggered relative to one another in separate planes and the salicylate ions stack above the heterodimer hydrogen bonded rings in the shorter spacing and the hydrogen bonded rings stack above one another in the larger spacing with a cytosinium molecule above a salicylate molecule and vice versa. The presence of a fluorine atom does not disrupt the base-stacking. However, there is a fluorine-fluorine interaction between the green and blue layers (F...F distance of 2.710 Å) which could contribute to the large difference between the distances of the green / blue and yellow / green separations (Table 4.3).

Like the non-fluorinated molecular complex, there is a weak C-H...O hydrogen bond at the tips of the L shaped unit (Figure 4.11). This hydrogen bond is from a *para* hydrogen atom of the salicylate molecule to the carbonyl oxygen of a cytosinium molecule in the case of the cytosinium salicylate molecular complex. However in the 5-fluorocytosinium salicylate molecular complex, this hydrogen bond is from a meta aromatic hydrogen atom to an oxygen of the carboxylate group on the salicylate molecule. Thus both the molecules involved, and the relative orientation of this hydrogen bond, are different between these

two complexes. The C-H...O hydrogen bond distance is 3.462(2) Å, which is shorter than that found in the non-fluorinated complex (3.527(3) Å).

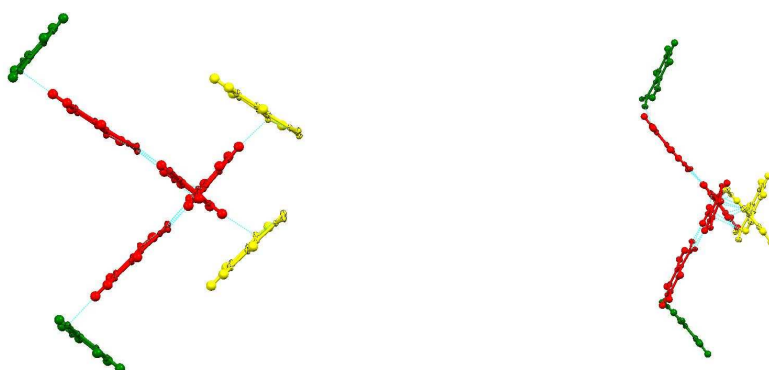


Figure 4.11. Molecules hydrogen bonded to the tips of the L-shape created by the pairs of left, cytosinium salicylate and right, 5-fluorocytosinium salicylate molecular complexes.

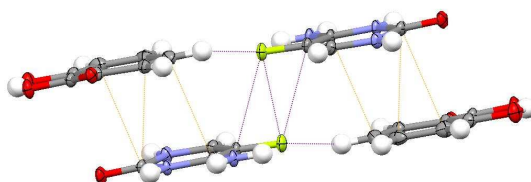


Figure 4.12. Base-stacking interactions and fluorine interactions in the molecular complex of 5-fluorocytosinium salicylate. The fluorine interactions are represented by purple dashed lines and the base stacking interactions by luminous orange dashed lines.

The sole difference between the crystallisation components is the fluorination of the cytosine molecule. Therefore the presence of this fluorine atom could be considered to be the driving force behind the differences between the crystal structures of the two. There are two fluorine interactions in this molecular complex (Figure 4.12). The first interaction is a fluorine-fluorine interaction between two separate layers in addition to the base stacking. This is between the shortest spaced layers. The fluorine interaction between the layers is from one 5-fluorocytosinium molecule to another. This interaction is of F...F length 2.744(2) Å which is within the sum of the van der Waals radii for two fluorine atoms of 2.94 Å. The second fluorine interaction is a C-H...F weak hydrogen bond between a salicylate molecule and a 5-fluorocytosinium molecule within the same layer. This has a C...F distance of 3.278(2) Å. This interaction forms in preference to the C-H...O hydrogen bond found in the cytosinium salicylate molecular complex (Figure 4.13). The differences between the structures can be attributed to the presence of the fluorine atom and explain why these two complexes are not isostructural.

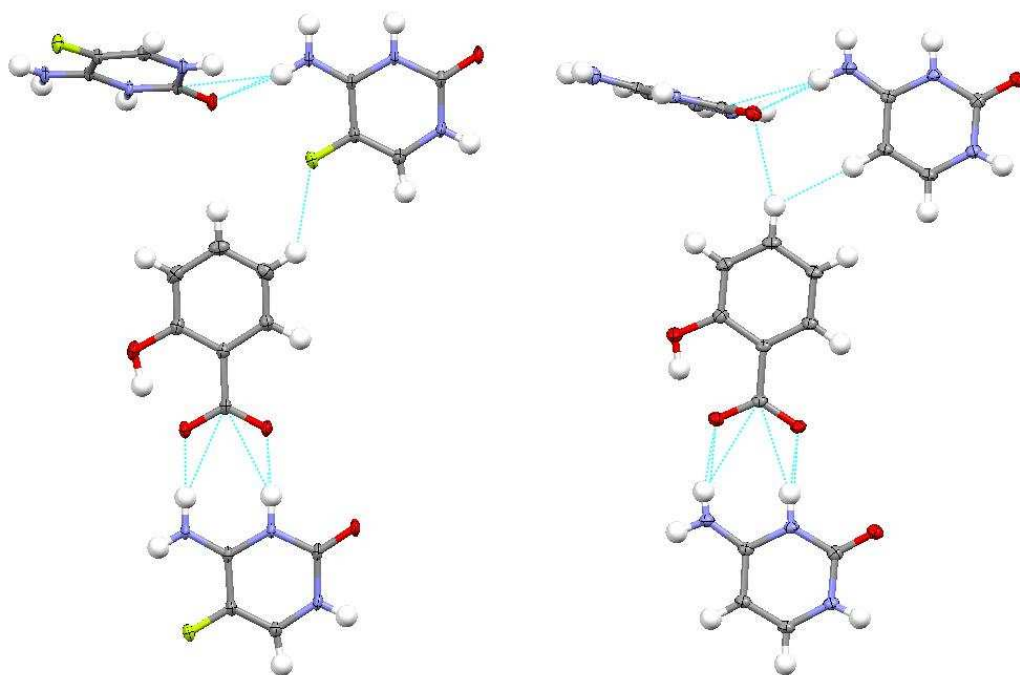


Figure 4.13. Influence of the fluorine atom in the hydrogen bonding in the related structures of 5-fluorocytosinium salicylate (left) and cytosinium salicylate (right).

4.3. 1:1 Molecular Complex of Cytosinium 5-Fluorosalicylate

Cytosine and 5-fluorosalicyclic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.4mm x 0.4mm x 0.5mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. The molecular complex forms in a 1:1 ratio and crystallographic data are given in Table 4.2.

A hydrogen atom is transferred from the 5-fluorosalicyclic acid molecule to the cytosine molecule creating two charged species in the structure. A hydrogen-bonded heterodimer is formed between the resultant cytosinium and 5-fluorosalicylate molecules as in the cytosinium and 5-fluorocytosinium salicylate molecular complexes (Figure 4.14).

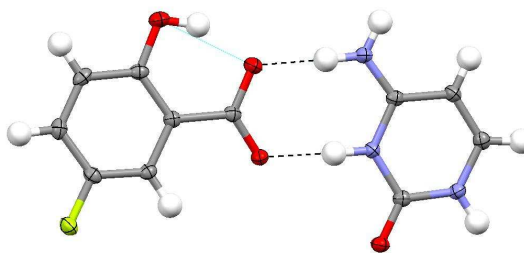


Figure 4.14. Hydrogen bonded heterodimer formed in the cytosinium 5-fluorosalicylate molecular complex.

The two N-H...O hydrogen bonds that make up this hydrogen-bonded ring are of moderate strength. The first is between the amine group of the cytosinium and the carboxylate group and has an N...O distance of 2.760(2) Å. The second has an N...O distance of 2.785(2) Å and is from a protonated ring nitrogen on the cytosinium to the carboxylate. There is also an intramolecular O-H...O hydrogen bond which helps to stabilise the charge on the 5-fluorosalicylate molecule which is moderately strong with an O...O distance of 2.549(2) Å. The 5-fluorosalicylate ions are parallel to one another in different planes and the cytosinium molecules are also parallel to one another. However, the hydrogen bonded unit is not planar and so the planes generated by cytosinium molecules are not parallel to those generated by 5-fluorosalicylate molecules.

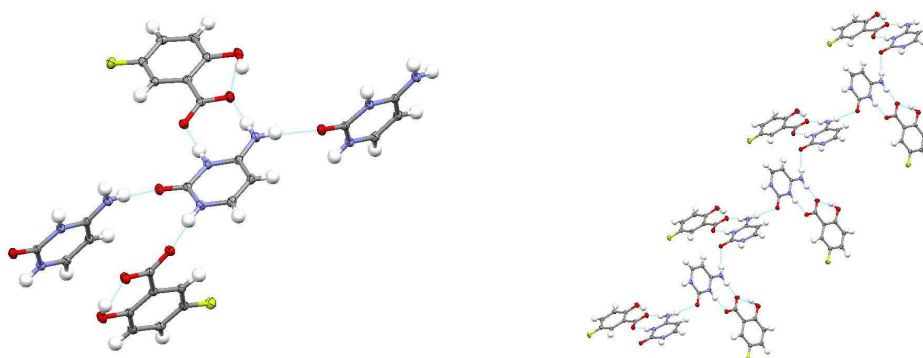


Figure 4.15. Left, the hydrogen bonding environment of the cytosinium molecule in the cytosinium 5-fluorosalicylate molecular complex; right, chain of hydrogen-bonded heterodimers.

The cytosinium molecule is involved in a total of five hydrogen bonds: two make up the hydrogen bonded ring, one is formed with a 5-fluorosalicylate molecule, and two are formed with other cytosinium molecules (Figure 4.15). The hydrogen bonds between cytosinium molecules are equivalent and between the amine groups and the carbonyl oxygen atoms with the cytosinium molecules aligned approximately perpendicular to one another as before. These hydrogen bonds are both of N...O distance of 3.040(2) Å and stepped chains of heterodimers are formed along the *ac* direction creating an L-shaped unit in the same manner as those formed in both the cytosinium and 5-fluorocytosinium

salicylate molecular complexes (Figure 4.15, right), with an angle of $\sim 112^\circ$ created at the centre of the L shape, intermediate between the other two related molecular complexes. The N \cdots O distance here is much longer than those found in the two related structures. This is due to different interactions in this region of the heterodimer; in the cytosinium salicylate and 5-fluorocytosinium salicylate molecular complexes, the intermolecular interactions around this hydrogen bond are similar and thus the N \cdots O distances are ~ 2.87 Å. However, in the cytosinium 5-fluorosaliclyate molecular complex, there is a C-H \cdots O weak hydrogen bond from the hydrogen in the 4- position of a 5-fluorosaliclyate molecule and the oxygen of one of the cytosine molecules making up the L shape (Figure 4.16). This hydrogen bond has a C \cdots O length of 3.425(2) Å. The second interaction consists of fluorine interactions with the ring of the cytosine molecule. These interactions combine to move the cytosine slightly out of position in comparison to the other two structures to give an extended N \cdots O distance.

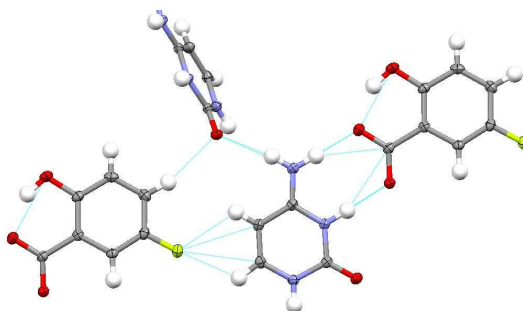


Figure 4.16. Fluorine interactions around the link between each arm of the L shape unit in the cytosinium 5-fluorosaliclyate molecular complex.

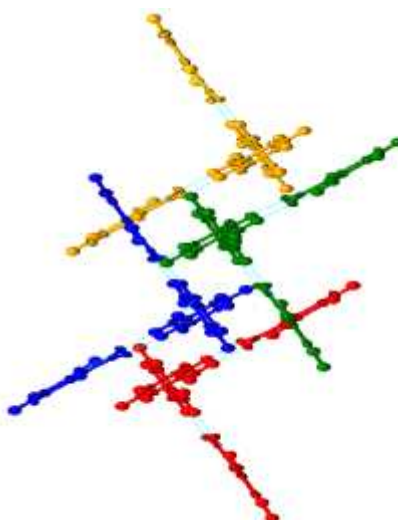


Figure 4.17. The herringbone structure of hydrogen bonded L shaped chains in the cytosinium 5-fluorosaliclyate molecular complex.

These chains are linked to one another through N-H \cdots O hydrogen bonds between a ring nitrogen and the carboxylate group on a neighbouring 5-fluorosaliclyate molecule. These

hydrogen bonds are of moderate strength and of N...O distance of 2.722(2) Å. This creates a herring-bone structure (Figure 4.17).

Figure 4.18 clearly shows the presence of base-stacking in this molecular complex. This interaction has been present in all of the related structures with the bases in very similar orientation relative to one another. From here it can be seen that there is base stacking between a red chain and a blue chain, a green chain and another blue chain and a yellow and green chain. There are two layer spacings and these are approximately 3.139 and 3.121 Å (Table 4.3). There is only base stacking interactions present between the yellow / green layer and the blue / red layer in the structure due to the orientation of the different layers. Figure 4.19 exemplifies this, with the nearest molecules to the original pair being perpendicular to the original. This means there cannot be any stacking interaction between green and blue. This means that there are only two unique stacking interactions present between the layers and the blue and green layers present in Figure 4.17 show this.

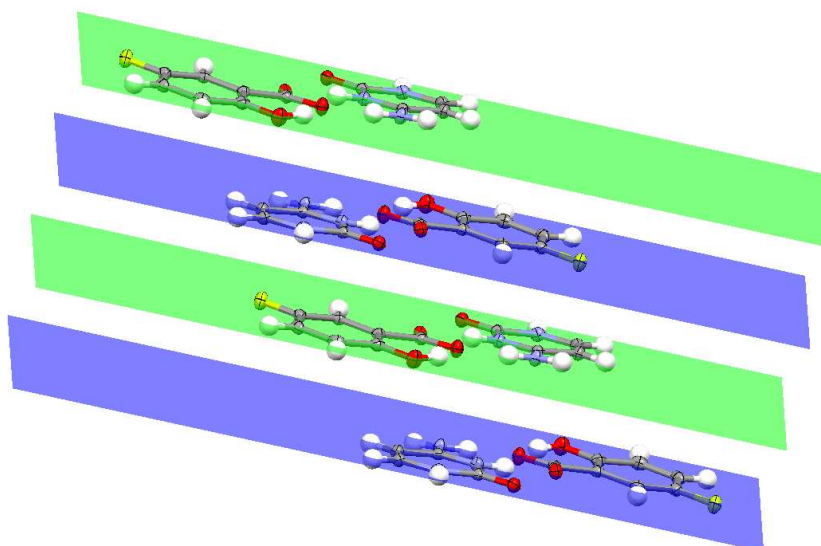


Figure 4.18. Base-stacking in the cytosinium 5-fluorosalicylate molecular complex.

Table 4.3 Comparison of the spacing between layers in the cytosinium salicylate, 5-fluorosalicylate and cytosinium 5-fluorosalicylate molecular complexes.

Layer Spacings (Å)	Cytosinium salicylate	5-Fluorocytosinium salicylate	Cytosinium 5-fluorosalicylate
Blue / Green	3.164	3.208	3.139
Green / Blue	3.136	3.150	3.121

Like the two other related structures, there is a core part of the structure that is very similar and the changes in the more extended structure are due to the presence of the fluorine substituent on one of the components. At the tips of the L shaped unit, there are two C-H...O hydrogen bonds between the para hydrogen atom of a 5-fluorosalicylate molecule and the oxygen of a cytosinium molecule. The arrangement of these two elements relative to each other is almost identical to that found in the cytosinium salicylate molecular

complex. The presence of the fluorine atom therefore does not disrupt this interaction. This hydrogen bond has an C...O distance of 3.475(2) Å (Figure 4.19).

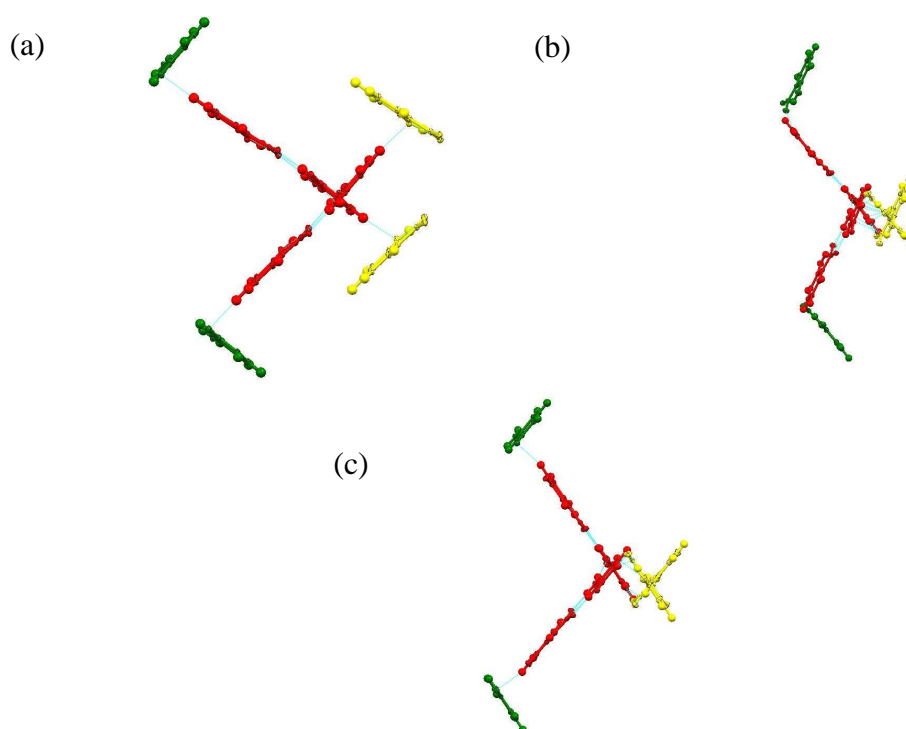


Figure 4.19. Molecules hydrogen bonded to the tips of the L-shape created by the pairs in the molecular complexes of (a), cytosinium salicylate, (b) 5-fluorocytosinium salicylate and (c) cytosinium 5-fluorosalicylate.

Table 4.4 Angle created at corner of L shape created in the three related structures.

Molecular Complex	Angle between the pairs making the L shape (°)
Cytosinium salicylate	90
5-fluorocytosinium salicylate	117
Cytosinium 5-fluorosalicylate	112

However, the interactions at the apex of these L-shaped units are quite different. The fluorine of the 5-fluorosalicylate molecule forms a DDHHA bifurcated weak C-H...F hydrogen bond with an aromatic hydrogen on a cytosinium molecule in the same plane. The C...F distances are 3.027(2) for the major component and 3.078(3) Å for the minor component (Figure 4.20). The other interaction in which the fluorine is involved is an F...O close contact with the hydroxyl group of a 5-fluorosalicylate molecule which lies approximately perpendicular to it. This interaction has an F...O length of 2.925(2) Å and is within the sum of the van der Waals radii of 2.99 Å for fluorine and oxygen.

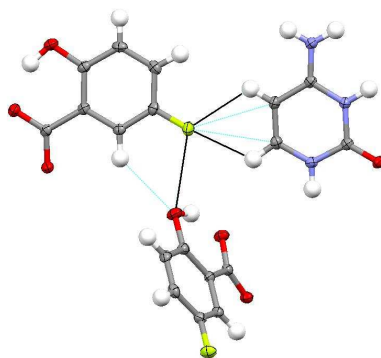


Figure 4.20. Fluorine interactions in the molecular complex of cytosinium 5-fluorosalicylate.

4.4. 2:1:2 Molecular Complex of 5-Fluorocytosine and 5-Fluorosalicylic Acid Hydrate

5-Fluorocytosine and 5-fluorosalicylic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.2mm x 0.2mm x 0.2mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. The crystallographic data are given in Table 4.2.

The molecular complex forms in a 2:1:2 ratio of 5-fluorocytosine to 5-fluorosalicylic acid to water. There are 5 molecules in the asymmetric unit, and the crystal belongs to the space group P-1. Whilst the ratio in the product is presented as 2:1:2, one cytosine has gained a hydrogen from the carboxyl group of the 5-fluorosalicylic acid essentially creating four different molecular species, two of which carry a charge: a neutral 5-fluorocytosine and a protonated 5-fluorocytosine (5-fluorocytosinium ion), a deprotonated 5-fluorosalicylic acid (5-fluorosalicylate), and two water molecules (Figure 4.21, left).

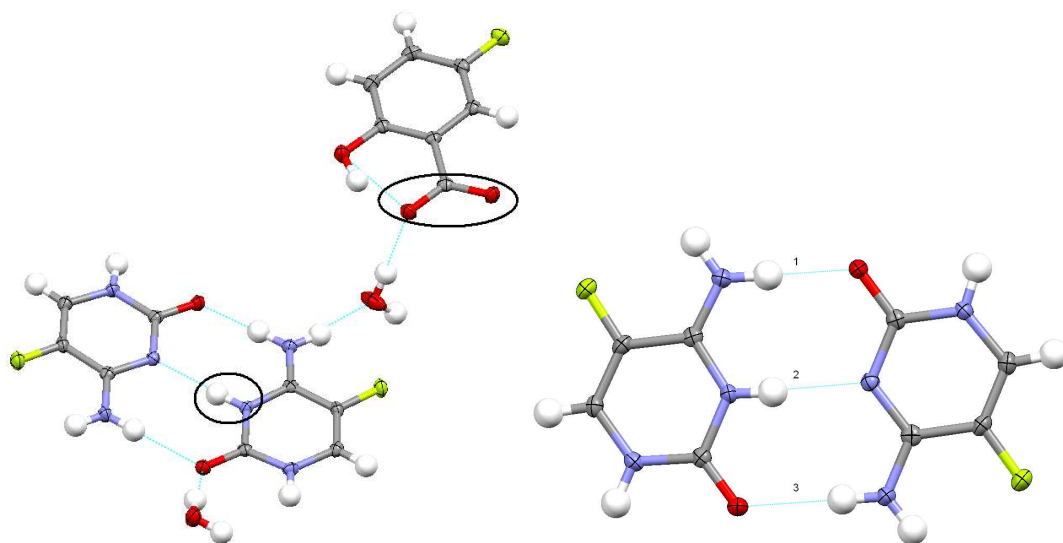


Figure 4.21. Left, The asymmetric unit of the molecular complex showing proton transfer from one 5-fluorosalicylic molecule to one 5-fluorocytosine molecule; right the pseudo Watson-Crick base pair.

The proton transfer allows the formation of a pseudo-Watson-Crick base pairing dimer between the neutral and protonated 5-fluorocytosine molecules. This pseudo-Watson-Crick base pair has three hydrogen bonds forming the supramolecular unit, one of which is charge assisted (Figure 4.21, right). This grouping is not possible for purely neutral 5-fluorocytosine molecules. The hydrogen bonds in this 5-fluorocytosine base pair are all of moderate strength. However, the three hydrogen bonds (Table 4.5) are shorter than those found in the molecular complex of ethyl-guanine and cytosine (EGMCYT10 [131]). This is a consequence of the central hydrogen bond being charge assisted. The hydrogen bond lengths are consistent with all the other molecular complexes reported in this thesis that show this primary hydrogen bonding motif.

Table 4.5. Hydrogen bond distances of the pseudo-Watson-Crick hydrogen bonding motif in the 5-fluorocytosine and 5-fluorosalicic acid hydrate molecular complex. The numbering refers to that shown in Figure 4.21.

Hydrogen Bond	D...A Distance
1. N-H---O	2.740(2) Å
2. N---H-N	2.837(2) Å
3. O---H-N	2.967(2) Å

Two base pair motifs are linked via the type B hydrogen bonding motif base-pairing. This base pair motif is between two cytosine molecules employing the same area of each cytosine. The two hydrogen bonds are equivalent N-H...O hydrogen bonds with an inversion centre located in the centre of this type B base pair and are of moderate strength with an N...O distance of 2.774(2) Å. This creates a four molecule 5-fluorocytosine 5-fluorocytosinium unit (Figure 4.22, left). At this point, the hydrogen bonded chain is terminated through hydrogen bonding to 5-fluorosalicylate and water molecules (Figure 4.22, right).

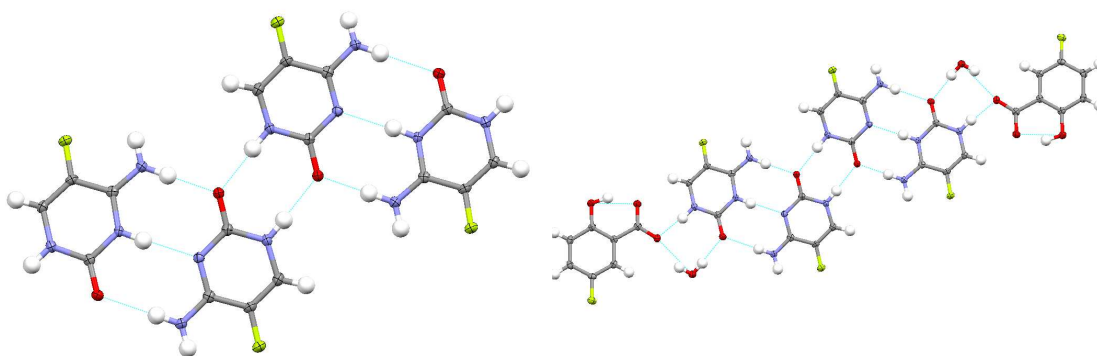


Figure 4.22. Left, the pseudo-Watson-Crick and type B hydrogen bonding motifs and right, the termination of the chain in the 5-fluorocytosine and 5-fluorosalicic acid hydrate molecular complex to create an eight molecule unit.

This chain is terminated at both ends through the combination of a 5-fluorosalicylate and a water molecule (Figure 4.22, right) to create an eight molecule unit. The 5-fluorosalicylate molecule forms two hydrogen bonds to a 5-fluorocytosinium molecule, one moderate and one weak. The first of these is an N-H...O hydrogen bond from the protonated nitrogen of the cytosine ring, to one of the oxygens of the carboxylate group of the 5-fluorosalicylate molecule. This hydrogen bond is of moderate strength and is of N...O distance of 2.684(2) Å. The second hydrogen bond is a weak C-H...O hydrogen bond between an aromatic carbon on the 5-fluorocytosinium molecule and the other oxygen of the carboxylate group of the 5-fluorosalicylate molecule. This has a C...O distance of 3.097(2) Å. One of the water molecules also helps terminate this chain by being involved in two hydrogen bonds within the same plane. The first of these O-H...O hydrogen bonds is between the water and the carboxylate group and is of moderate strength with an O...O distance of 2.846(2) Å. The other hydrogen bond is between the water and the carbonyl oxygen of the nearby cytosine. This hydrogen bond is also of moderate strength and is of O...O distance of 2.857(2) Å.

On each side of the supramolecular unit, there are two additional hydrogen bonds to water molecules. Both of these hydrogen bonds involve the amine groups of the 5-fluorocytosine and 5-fluorocytosinium molecules to the oxygens of the two independent water molecules. The first is from a 5-fluorocytosinium molecule to a water molecule lying within the same layer (yellow in Figure 4.23) and this connects supramolecular units together within the same plane. These hydrogen bonds are of moderate strength and of N...O distance of 2.885(2) Å. These hydrogen bonds in combination with two 5-fluorocytosine 5-fluorocytosinium four molecule units, form six molecule $R_4^6(12)$ hydrogen bonded rings (Figure 4.24). The second hydrogen bond is from a 5-fluorocytosine molecule to a water molecule and this connects supramolecular units together in parallel layers (green in Figure 4.23). These are also moderate hydrogen bonds with N...O distance of 2.793(2) Å.

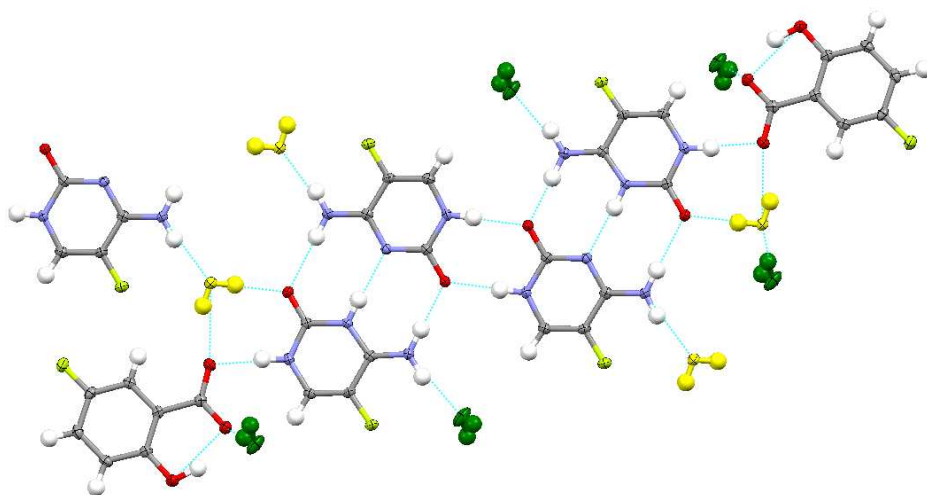


Figure 4.23. The hydrogen bonds involving the four molecule 5-fluorocytosine 5-fluorocytosinium unit in the 5-fluorocytosine 5-fluorosalicic acid hydrate molecular complex. The water molecules which terminate the cytosine chains are shown in yellow and those that connect planes in different layers are in green.

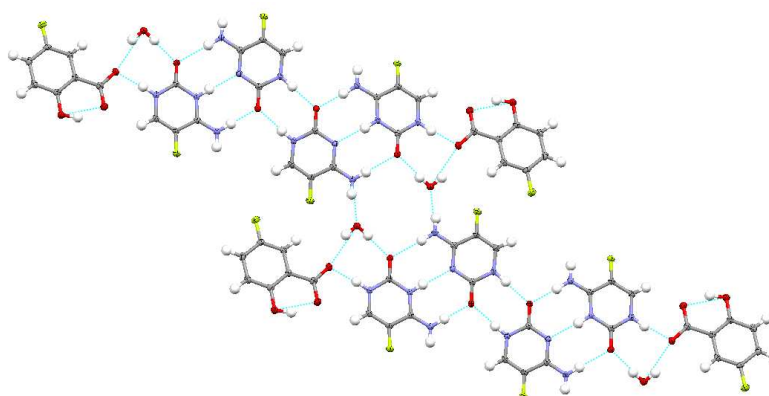


Figure 4.24 The six molecule hydrogen bonded ring in the 5-fluorocytosine 5-fluorosalicic acid hydrate molecular complex connecting two eight molecule units together.

This second water molecule sits approximately perpendicular to the original chain and this forms hydrogen bonds to two different molecules in different layers, one to a 5-fluorosalicylate molecule and one to a water molecule (green in Figure 4.23, Figure 4.25). The first hydrogen bond is to one of the oxygen atoms of a carboxylate group. This hydrogen bond has an O...O distance of 2.817(2) Å and is of moderate strength. The other hydrogen bond is from the water to another independent water molecule (yellow in Figure 27). This O-H...O hydrogen bond length is 2.992(2) Å and it is of moderate strength.

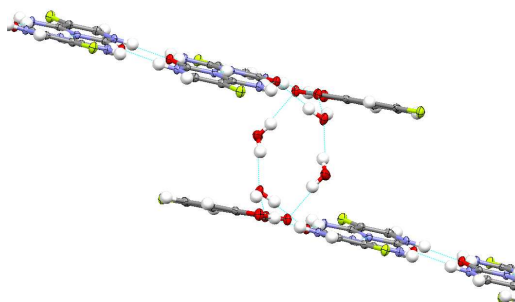


Figure 4.25. Hydrogen bonds involving the water molecule which connects the layers together in the 5-fluorocytosine 5-fluorosalicyclic acid hydrate molecular complex.

The presence of fluorine does not disrupt the base-pairing motif but there are some significant fluorine interactions in the structure. There is a fluorine-fluorine interaction between 5-fluorocytosinium molecules in different layers (Figure 4.26, left). This interaction is of $F\cdots F$ length of $2.787(2)$ Å which is shorter than the sum of the van der Waals radii for two fluorine atoms of 2.94 Å. This acts as a link between layers in combination with base-stacking effects and the hydrogen bonding involving the water molecule. There is also a weak $C-H\cdots F$ hydrogen bond (Figure 4.26, right) from the ortho hydrogen on a 5-fluorosalicylate molecule to the fluorine on one of the 5-fluorocytosine molecules. This is in the same plane and acts as a link between separate units. This hydrogen bond has a $C\cdots F$ length of $3.263(2)$ Å.

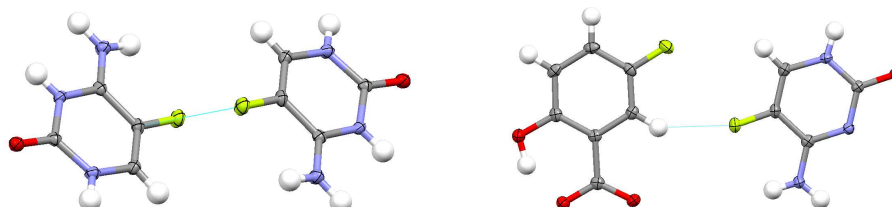


Figure 4.26. Fluorine interactions in the 5-fluorocytosine 5-fluorosalicyclic acid hydrate molecular complex. Left, fluorine-fluorine interaction between 5-fluorocytosinium molecules; right, weak $C-H\cdots F$ hydrogen bond.

There are several base-stacking effects in this structure (Figure 4.27). The first of these takes the form of a cytosine pair stacked on top of another cytosine pair. These are not identical as the top pair has been rotated 180° about the pseudo-Watson-Crick hydrogen bonding of the base pair. This results in the 5-fluorocytosinium molecules stacking upon 5-fluorocytosine molecules and vice versa. The distance between the two based paired planes is approximately 3.183 Å which is in line with that seen for other structures with the pseudo-Watson-Crick bonding motif. The second base stacking interaction is between the 5-fluorosalicylate molecules and the cytosine base pairs. This distance is approximately 3.077 Å.

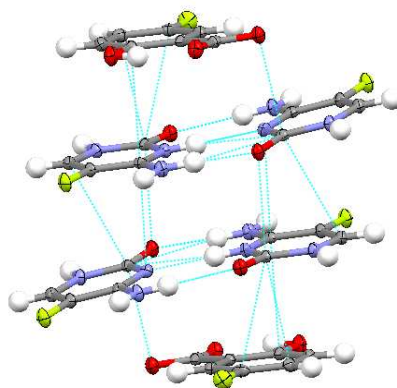


Figure 4.27. Stacking interactions in the 5-fluorocytosine 5-fluorosalicic acid hydrate molecular complex

4.5. 2:1:2 Molecular Complex of Cytosine and 5-Fluorosalicic Acid Hydrate

Cytosine and 5-fluorosalicic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.4mm x 0.4mm x 0.4mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are given in Table 4.2.

The molecular complex forms in a 2:1:2 ratio of cytosine to 5-fluorosalicic acid to water and is isostructural with the 2:1:2 5-fluorocytosine 5-fluorosalicic acid molecular complex. There are 5 molecules in the asymmetric unit, and the crystal belongs to the space group P-1. Whilst the ratio in the product is presented as 2:1:2, one cytosine has gained a hydrogen from the carboxyl group of the 5-fluorosalicic acid essentially creating four different molecular species, two of which carry a charge; a neutral cytosine and a protonated cytosine (cytosinium); a deprotonated 5-fluorosalicic acid (5-fluorosalicylate); and two water molecules (Figure 4.28, left). The proton transfer allows the formation of a pseudo-Watson-Crick base pairing dimer between the neutral and protonated cytosine molecules. This pseudo-Watson-Crick base pair has three hydrogen bonds forming the supramolecular unit, one of which is charge assisted (Figure 4.28, right).

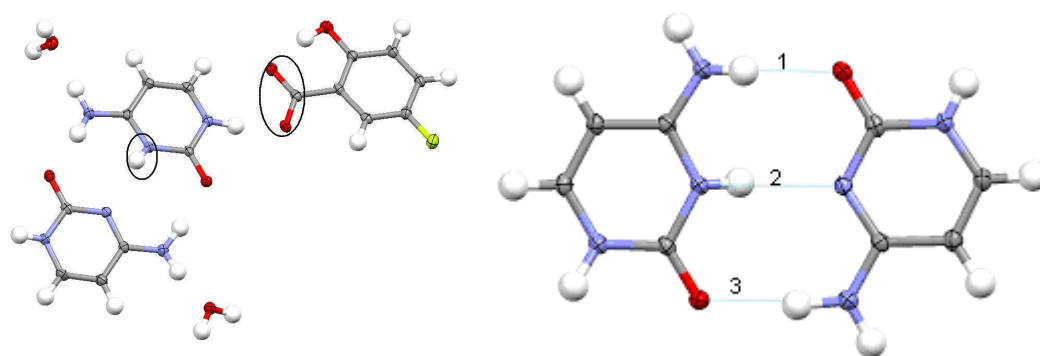


Figure 4.28. Left, the asymmetric unit of the molecular complex of cytosine 5-fluorosalicic acid hydrate showing proton transfer from the 5-fluorosalicic acid molecule to one cytosine molecule; right, the pseudo Watson-Crick base pair.

This structure is very similar in its construction to the 2:1:2 5-fluorocytosine and 5-fluorosalicic acid hydrate molecular complex. The hydrogen bonds making up the pseudo-Watson-Crick unit are all of moderate strength and are again shorter than those found in the molecular complex of ethyl-guanine and cytosine (EGMCYT10). The hydrogen bond distances (Table 4.6) are very similar to those found in the 5-fluorosalicic acid and 5-fluorocytosine hydrate molecular complex. The asymmetry in the N-H...O hydrogen bonds is also common to other pseudo-Watson-Crick units reported in this thesis. The majority of structures show the hydrogen bond between the amine group of the neutral cytosine and the oxygen atom of the charged cytosinium molecule to be longer than the equivalent hydrogen bond involving the cytosinium amine group.

Table 4.6 Hydrogen bond distances of the pseudo-Watson-Crick hydrogen bonding motif in the cytosine 5-fluorosalicic acid hydrate molecular complex

Hydrogen Bond	D...A Distance
1. N-H---O	2.767(2) Å
2. N---H-N	2.852(2) Å
3. O---H-N	2.943(2) Å

Much like the 5-fluorocytosine 5-fluorosalicic acid hydrate molecular complex, two pseudo Watson Crick base pair motifs are linked via type B hydrogen bonding motif base-pairing. This base pair motif is between two cytosine molecules employing the same area of each cytosine. The two hydrogen bonds are equivalent N-H...O hydrogen bonds with an inversion centre located in the centre of this type B base pair. They are of moderate strength and have a N...O distances of 2.788(2) Å. This creates a four molecule cytosine unit (Figure 4.29, left). At this point, the chain is terminated through 5-fluorosalicylate and water molecules (Figure 4.29, right).

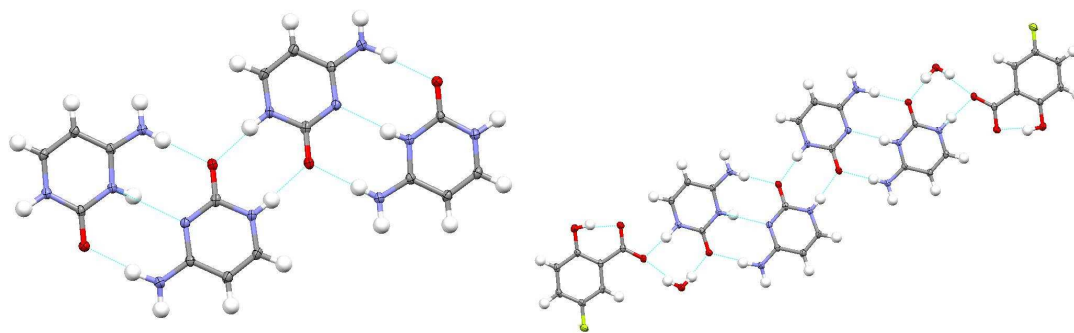


Figure 4.29. Left, link between two pseudo-Watson Crick base pairs via type B hydrogen bonding motif; right, termination of the chain in the cytosine and 5-fluorosalicylic acid hydrate molecular complex.

One weak and one moderate strength hydrogen bond is created between a cytosinium molecule and a 5-fluorosalicylate acid molecule. The first of these hydrogen bonds is a moderate strength N-H...O hydrogen bond from a protonated ring nitrogen of the cytosinium, to one of the oxygen atoms of the carboxylate group of the 5-fluorosalicylate molecule. This hydrogen bond has an N...O distance of 2.684(2) Å. The weak hydrogen bond is a C-H...O hydrogen bond between an aromatic carbon on the cytosinium and the other oxygen of the carboxylate group of the 5-fluorosalicylate molecule. The C...O distance is 3.185(3) Å. A water molecule combines with this 5-fluorosalicylic acid molecule to terminate the chain via two hydrogen bonds within the same plane (green in Figure 4.30). The first of these O-H...O hydrogen bonds is between the water and the carboxylate group and is of moderate strength and O...O distance of 2.749(2) Å. The other hydrogen bond is between the water and the carbonyl oxygen of the nearby cytosine. This hydrogen bond is also of moderate strength and have an O...O distance of 2.838(2) Å.

Similarly to the 5-fluorocytosine 5-fluorosalicylic acid hydrate molecular complex, on each side of the supramolecular unit, there are two additional hydrogen bonds to water molecules. These two hydrogen bonds involve the amine groups of the cytosine and cytosinium molecules to the oxygen atom of the two independent water molecules. The first is to a water molecule lying within the same layer (green in Figure 4.30) and these molecules connect supramolecular units together within the same plane. These hydrogen bonds are of moderate strength and have an N...O distance of 2.879(2) Å. This water molecule, utilising all three of its hydrogen bonds, connects cytosine cytosinium four molecule units together through a six molecule hydrogen bonded ring in the same manner as that found in the 5-fluorocytosine 5-fluorosalicylic acid hydrate molecular complex (Figure 4.31). The second of these two hydrogen bonds is to a water molecule which connects supramolecular units together in parallel layers (purple in Figure 4.30). These are also of moderate strength and have an N...O distance of 2.808(2) Å.

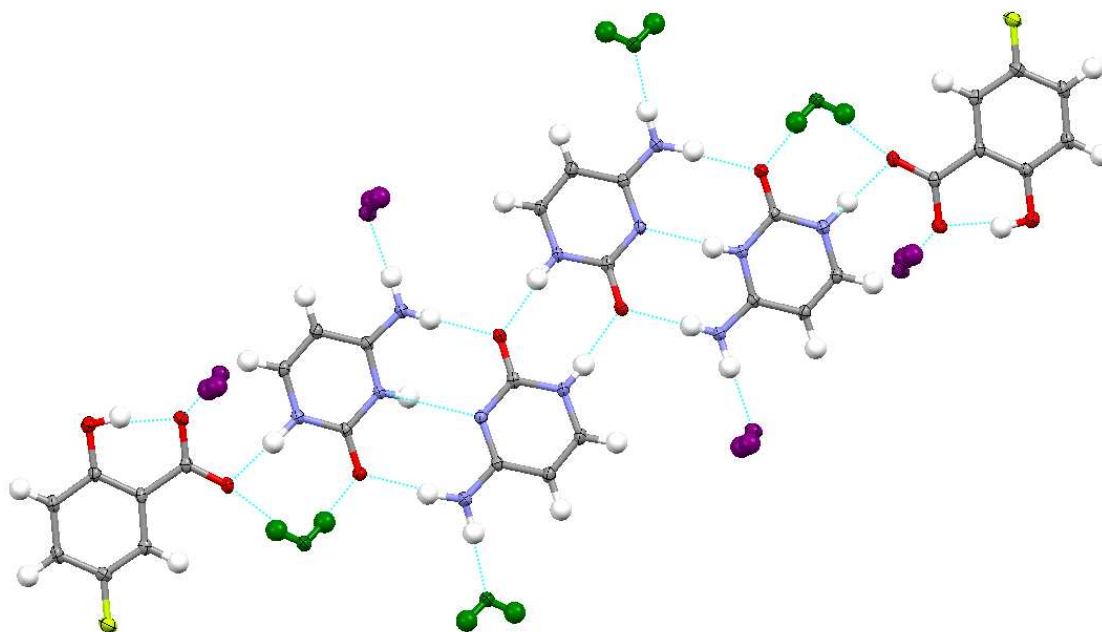


Figure 4.30. The hydrogen bonds involving the four molecule cytosine cytosinium unit in the cytosine 5-fluorosalicylic acid hydrate molecular complex. The water molecules which terminate the cytosine chains are shown in green and those that connect planes in different layers are in purple.

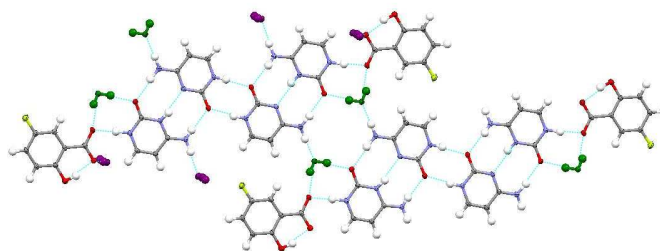


Figure 4.31. Six molecule hydrogen bonded ring in the cytosine 5-fluorosalicylic acid hydrate molecular complex.

This water molecule sits approximately perpendicular to the original chain and makes hydrogen bonds to two different molecules in different layers (purple in Figure 4.30, Figure 4.32), one to a 5-fluorosalicylate molecule and one to an independent water molecule. The first hydrogen bond is to one of the oxygen atoms of a carboxylate group. This hydrogen bond has an O...O distance of 2.799(2) Å and is of moderate strength. The other hydrogen bond is from the water to another independent water molecule (green in Figure 4.30). This O...O hydrogen bond length is 2.892(2) Å and it is of moderate strength.

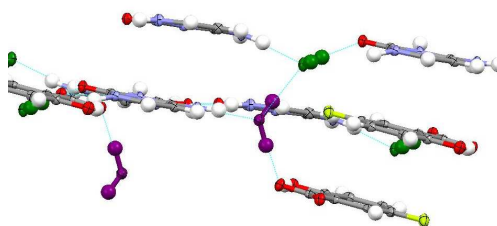


Figure 4.32. Hydrogen bonds involving the water molecule connecting the layers in the cytosine 5-fluorosalicylic acid hydrate molecular complex.

The presence of fluorine does not disrupt the base-pairing motif in the structure and the fluorine atom on the 5-fluorosalicylate molecule only takes part in one interaction. This is a weak hydrogen bond between the cytosine and 5-fluorosalicylate molecules with a C...F length of 3.584(2) Å (Figure 4.33).

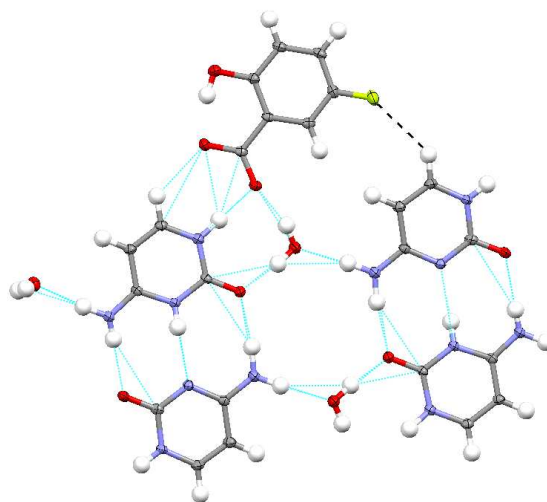


Figure 4.33. Weak C-H...F hydrogen bonds represented by a black dashed line, in the cytosine 5-fluorosalicylic acid hydrate molecular complex.

Base stacking is once again present. Pseudo-Watson-Crick base pairs stack upon one another (C in Figure 4.34), there is stacking between the base pair and a 5-fluorosalicylate (B and D in Figure 4.34) and also stacking between two 5-fluorosalicylate molecules (A in Figure 4.34). The layers in this structure are not as planar as in the other structures. Approximate distances between the layers are given in Table 4.7. These are similar to the spacings found in the related structures.

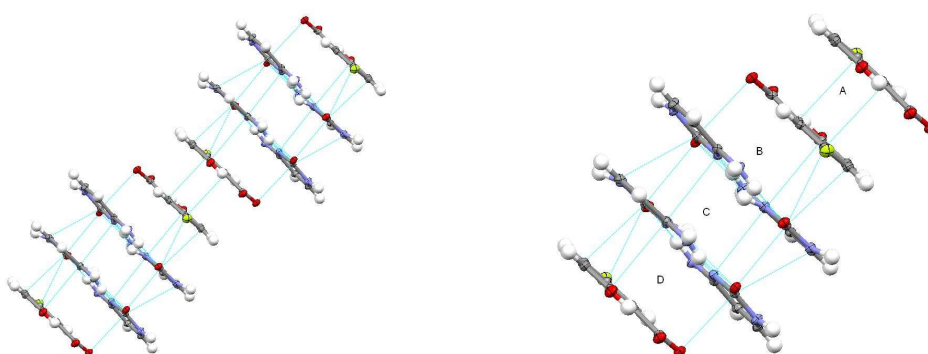


Figure 4.34. Stacking interactions present in the cytosine 5-fluorosalicylic acid hydrate molecular complex.

Table 4.7. Distances between planes in the molecular complex of cytosine 5-fluorosalicylic acid hydrate.

Base Stacking	Distances between the planes Å
A	3.275
B	3.310
C	3.161
D	3.247

This structure is isostructural with the fully fluorinated equivalent. The only difference between the two structures is the presence of a fluorine atom on the cytosine molecule in one structure that is not present in the other. It must be the interactions in which this fluorine molecule is involved that causes the two structures to differ slightly (Figure 4.35). The fluorine atom has interactions with molecules in the planes above one another and with molecules in the same plane in the 5-fluorocytosine 5-fluorosalicyclic acid hydrate molecular complex. However there are no fluorine-fluorine interactions between the planes in the cytosine 5-fluorosalicyclic acid hydrate molecular complex. The fluorine plays the important role in causing very subtle differences between these related structures showing the power of the addition of a fluorine atom in changes in a structure.

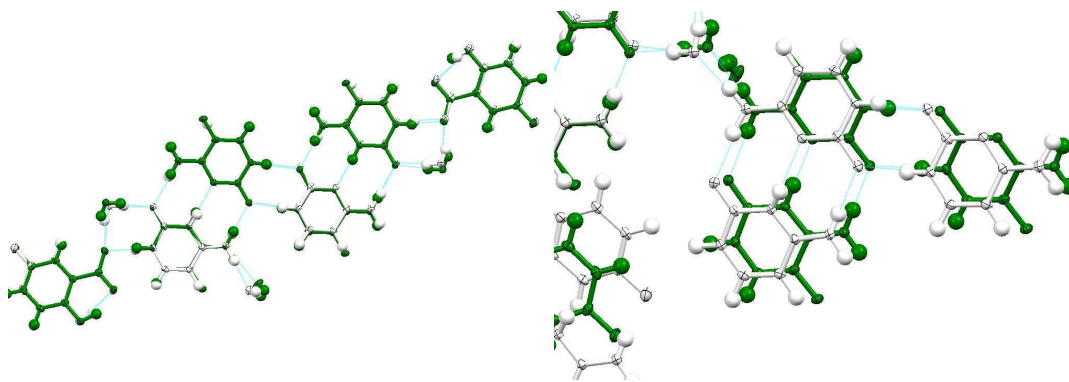


Figure 4.35. Left, overlay of the structures of cytosine 5-fluorosalicyclic acid hydrate (white) and 5-fluorocytosine and 5-fluorosalicyclic acid hydrate (green) and right, the two structures slowly moving out of sync.

4.6. 2:2:1 Molecular Complex of Cytosinium and 3-Hydroxybenzoate Hemihydrate

Cytosine and 3-hydroxybenzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.3mm x 0.3mm x 0.5mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 4.9.

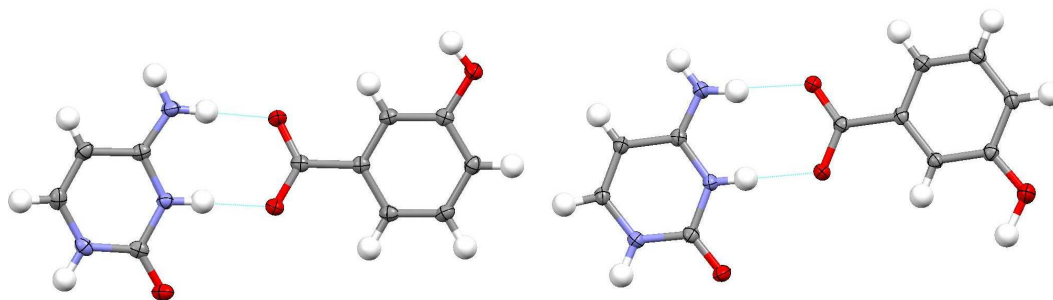


Figure 4.36. The two independent heterodimers in the molecular complex of cytosinium 3-hydroxybenzoate hemihydrate molecular complex.

The molecular complex forms in a 2:2:1 ratio of cytosine to 3-hydroxybenzoic acid to water. Both cytosine molecules have gained a hydrogen atom from the carboxyl groups of the two 3-hydroxybenzoic acid molecules and so this system shows 100% proton transfer. The asymmetric unit therefore has two charged cytosine molecules (cytosinium) and two deprotonated 3-hydroxybenzoic acid molecules (3-hydroxybenzoate). The primary hydrogen bonding motif formed is the hydrogen-bonded heterodimer, however, the two cytosinium molecules and two 3-hydroxybenzoic acid molecules form two independent hydrogen heterodimers within the asymmetric unit (Figure 4.36). Although the basic hydrogen bonding motif is the same, there are some noticeable differences. One of the main differences between the two dimers is the orientation of the carboxylate group relative to the plane of the 3-hydroxybenzoate molecule (Figure 4.37). The torsion of the carboxylate group is different in each case - for dimer 1 this is approximately 8.5° and for dimer 2 it is approximately 16° .

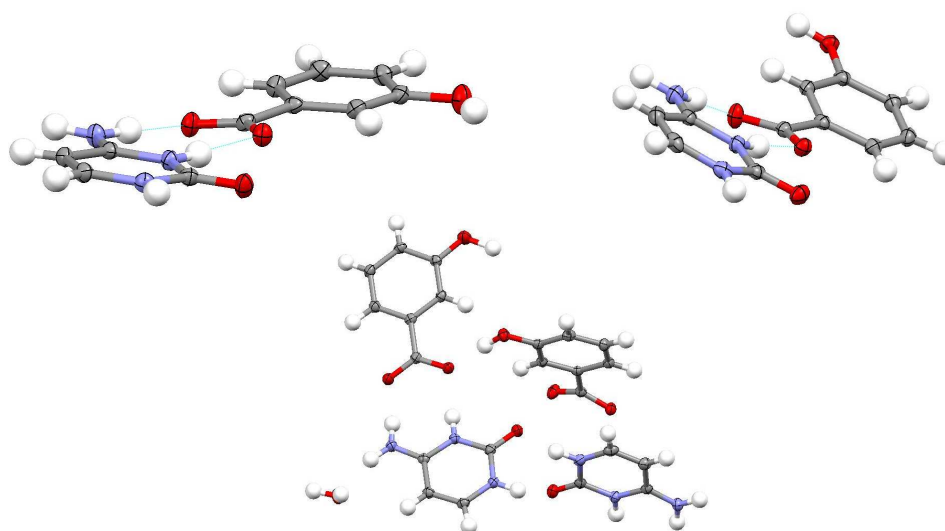


Figure 4.37. Top, the two independent hydrogen bonded heterodimers showing the increased torsion angle of the carboxylic acid group of one (top right) compared to the other (top left). Bottom, the asymmetric unit of the cytosinium 3-hydroxybenzoate molecular complex.

There are two N-H \cdots O hydrogen bonds comprising each heterodimer, and these are all of moderate strength. The hydrogen bond between the amine group and the carboxylate group has an N \cdots O distance of 2.773(2) Å in both dimers. The hydrogen bond between the ring

nitrogen and the carboxylate group has an N...O distance of 2.787(2) Å in the more planar dimer (dimer 1), and is considerably shorter in the more twisted dimer (dimer 2) having an N...O distance of 2.741(2) Å (Table 4.8).

Table 4.8. Hydrogen bonds of the primary bonding motif in dimer 1 and 2.

	N-H...O (amine to carboxylate oxygen) (Å)	N-H...O (protonated ring nitrogen to carboxylate oxygen) (Å)
Dimer 1	2.773(2)	2.787(2)
Dimer 2	2.773(2)	2.741(2)

Similarly to the cytosinium and 5-fluorocytosinium salicylate molecular complexes, each cytosinium molecule is involved in a total of five hydrogen bonds (Figure 4.38). The cytosinium in dimer 2 forms two hydrogen bonds to separate 3-hydroxybenzoate molecules, one as a hydrogen bond donor from the amine group to the carboxylate group of a salicylate molecule (N...O distance, 2.773(2) Å), and one as a hydrogen bond acceptor from a hydroxyl group of a salicylate molecule to the carbonyl group of the cytosinium (O...O distance, 2.719(2) Å). A third single hydrogen bond is formed from a ring nitrogen to the carbonyl of an independent cytosinium molecule (N...O distance, 2.741(2) Å) (Figure 4.38, right). The cytosinium in dimer 1 hydrogen bonds to a water molecule from the amine group through an N-H...O hydrogen bond (N...O distance, 2.780(2) Å), forms an N-H...O hydrogen bond from a ring nitrogen to a carboxylate group (N...O distance, 2.688(2) Å) and an O...H-N hydrogen bond between the carbonyl and a ring nitrogen on an independent cytosinium molecule (N...O, distance 2.760(2) Å) (Figure 4.38, left).

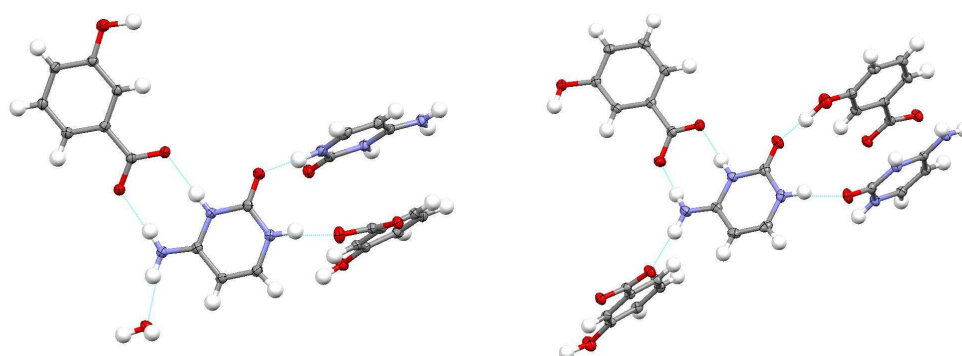


Figure 4.38. The two distinct hydrogen bonding environments of the independent cytosinium molecules in the cytosinium 3-hydroxybenzoate hemihydrate molecular complex.

A similar chain to that found in the salicylic acid molecular complex is formed between cytosinium 3-hydroxybenzoate heterodimers. However, in this case there are two independent dimers (dimers 1 and 2), which alternate along the chain (Figure 4.39). The links between dimers are of two types, an N-H...O hydrogen bond between a ring nitrogen on dimer 1 and the carbonyl group on dimer 2 with an N...O distance of 2.760(2) Å. The

second linker is an N-H...O hydrogen bond between an amine group on dimer 2 and a carboxylate oxygen on dimer 1 with an N...O distance of 2.790(2) Å.

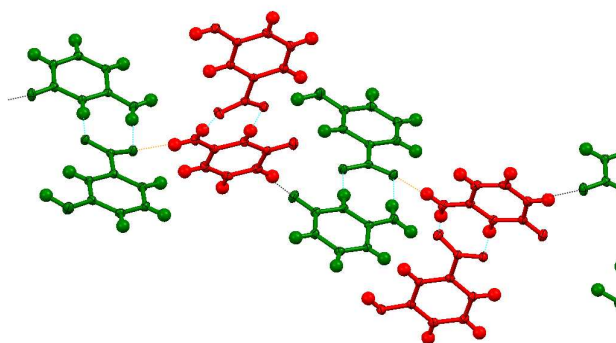


Figure 4.39. Alternating dimer 1 (red) and 2 (green) along an extended chain in the cytosinium 3-hydroxybenzoate hemihydrate molecular complex.

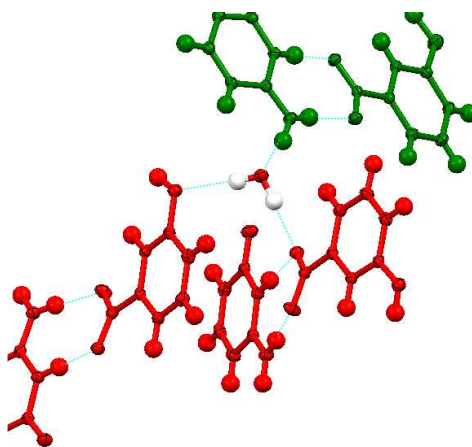


Figure 4.40. Water molecule acting as a link between three cytosinium and 3-hydroxybenzoate heterodimers.

The water molecule also links three separate heterodimers together via two single hydrogen bonds (Figure 4.40); the water acts as a hydrogen bond acceptor from dimer 1 and acts as a hydrogen bond donor to two dimers of type 2. The first of these hydrogen bonds is an N-H...O hydrogen bond from the hydrogen atom of the amine group of dimer 1 to the oxygen atom of the water molecule with an N...O distance of 2.780(2) Å. The second is an O-H...O hydrogen bond from the hydrogen of the water molecule to the hydroxy group of the 3-hydroxybenzoate, which is part of dimer 2 and has an O...O distance of 2.865(2) Å. The third is an O-H...O hydrogen bond to the carbonyl oxygen of a type 2 dimer with an O...O distance of 2.745(2) Å. The hydroxyl groups then play an important role in connecting the structure as they no longer participate in intramolecular hydrogen bonds. A 3-hydroxybenzoate molecule from a dimer 1 unit bridges these two dimer 2 units with a carboxylate oxygen acting as a hydrogen bond acceptor from the hydroxy group of the first and with the hydroxy group acting as a hydrogen bond donor to the carbonyl oxygen of a the cytosinium molecule in the second (Figure 4.41). These are

both O-H...O hydrogen bonds with O...O distances of 2.664(2) Å and 2.719(2) Å, respectively.

Chains of type 1 dimers are formed along the *c*-axis connected through weaker C-H...O hydrogen bonds from an aromatic hydrogen on a cytosinium molecule to the hydroxy group of a 3-hydroxybenzoate molecule in a neighbouring dimer. These hydrogen bonds have a C...O distance of 3.260(3) Å (Figure 4.42, left). Chains of dimer 2 are also formed along the *b*-axis but this time they are two units thick and held together by the bridging water molecules (Figure 4.42, right). Overall, the result of these hydrogen bonding patterns is the formation of layers of dimer 1 and dimer 2 alternating along the *a* axis (Figure 4.43). Only dimer 1 units are stacked upon other dimer 1 units and vice versa. Cytosinium molecules stack above 3-hydroxybenzoate molecules in both cases.

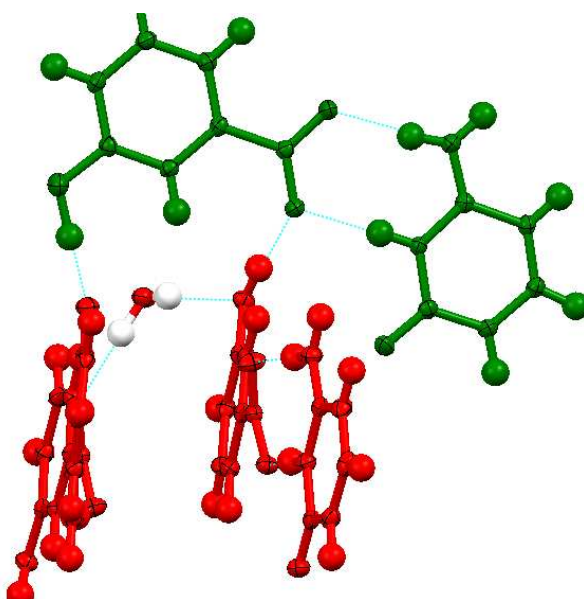


Figure 4.41. Water molecule acting as a link between three cytosinium and 3-hydroxybenzoate heterodimers.

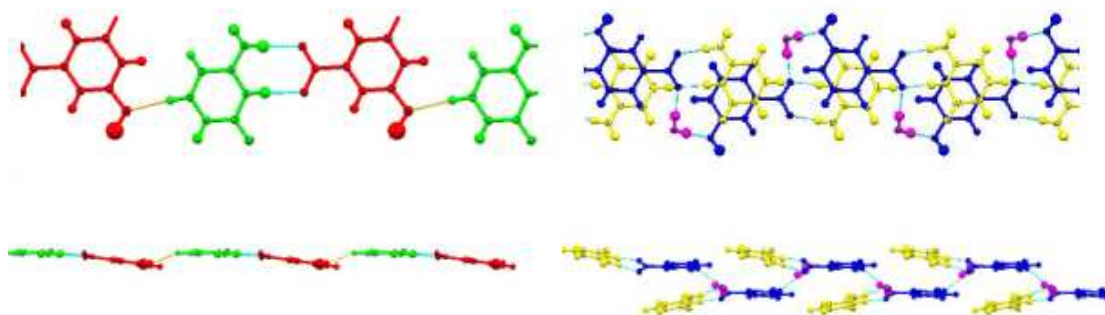


Figure 4.42. Chains formed with (a) just dimer 1 (red for cytosinium and green for the 3-hydroxybenzoate) and (b) just dimer 2 (yellow for cytosinium blue for the 3-hydroxybenzoate and purple for the water molecules) in the cytosinium and 3-hydroxybenzoate hemihydrate.

Figure 4.43 clearly shows that there are base stacking interactions in this structure as well. Here dimer one (red and green) stack on top of other dimer 1 units and the same can

be seen for the yellow and blue pair (dimer 2). When the red and green pair stacks on top of each other the cytosinium is stacked on top of the 3-hydroxybenzoate. The distance between these two is approximately 3.2 Å. The distance between the yellow and blue layers is approximately 3.140 Å. However, the layers are not totally planar and the distance between separate chains fluctuates as the chain is extended.

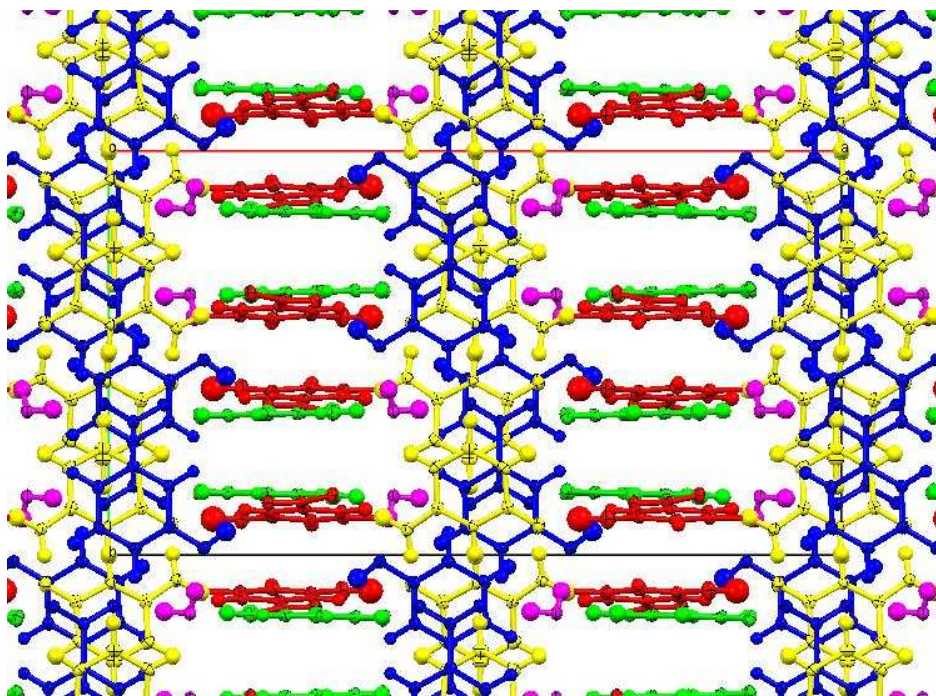


Figure 4.43. Packing in the cytosinium 3-hydroxybenzoate hemihydrate.

Table 4.9. Crystallographic data for the molecular complexes of cytosine and 5-fluorocytosine with 3 and 4-hydroxybenzoic acid.

Compound	Cytosinium 3-hydroxybenzoate hydrate	5-Fluorocytosine and 3-hydroxybenzoic acid	Cytosine and 4-hydroxybenzoic acid hydrate	5-Fluorocytosine and 4-hydroxybenzoic acid
Formula	C22 H24 N6 O9	C15 N6 H14 O5 F2	C22 H24 N6 O9	C11 N3 O4 H10 F1
Crystallisation Conditions	Methanol, Room Temperature	Methanol, Room Temperature	Methanol, Room Temperature	Methanol, Room Temperature
Molecular weight / gmol⁻¹	516.46	396.31	516.46	267.21
Temperature (K)	100 K	100 K	100 K	100 K
Space Group	Pbca	C2/c	P-1	P2 ₁ /c
<i>a</i> (Å)	24.8961(2)	19.9476(3)	9.9946(24)	6.8938(3)
<i>b</i> (Å)	13.7428(1)	6.7722(1)	10.6617(26)	25.1380(12)
<i>c</i> (Å)	13.5411(1)	26.0161(5)	15.1110(38)	9.1359(4)
α (°)	90	90	81.362(14)	90
β (°)	90	110.940(1)	71.277(13)	128.660(3)
γ (°)	90	90	87.191(14)	90
Volume (Å³)	4632.98(1)	3282.38(13)	1507.71(77)	1236.28(38)
<i>Z</i>	8	8	2	4
θ range (°)	2.3-40.7	1.7-27.8	1.4-29.6	1.6-27.5
Completeness	99.20%	98.60%	98.50%	98.50%
Reflections Collected	89837	43891	67140	31547
Independent	5320	3818	8343	2774
Refln (obs.I>2 σ (I))	4414	2611	6775	1661
<i>R</i>_{int}	0.0359	0.0676	0.0405	0.1104
Parameters	430	295	561	152
GooF on F²	1.035	1.003	1.035	1.167
<i>R</i>₁ (Observed)	0.036	0.0412	0.040	0.092
<i>R</i>₁ (all)	0.047	0.0761	0.052	0.165
<i>wR</i>₂ (all)	0.103	0.0982	0.120	0.210

4.7. 2:1 Molecular Complex of 5-Fluorocytosine and 3-Hydroxybenzoic Acid

5-Fluorocytosine and 3-hydroxybenzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.2mm x 0.2mm x 0.4mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 4.9.

The molecular complex forms in a 2:1 ratio of 5-fluorocytosine and 3-hydroxybenzoic acid and contains no hydrogen transfer with all molecules remaining neutral. This is in contrast to the molecular complex of cytosine with 3-hydroxybenzoic acid.

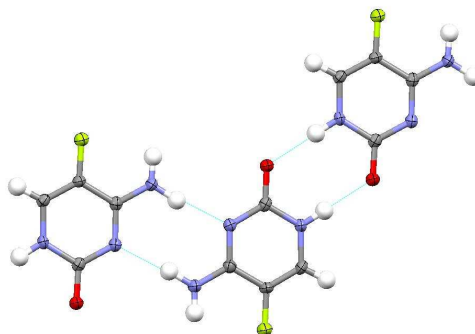


Figure 4.44. Combination of type A and B pseudo-base paired hydrogen bonding in the molecular complex of 5-fluorocytosine and 3-hydroxybenzoic acid.

The 5-fluorocytosine molecules form both type A and B hydrogen bonding motifs, both of which are base pairs consisting of two hydrogen bonds (Section 1.10). One of the base pairs has hydrogen bonds between the amine groups and unprotonated ring nitrogen atoms on neighbouring 5-fluorocytosine molecules (Figure 4.44). These two N-H...N hydrogen bonds are of moderate strength and have N...N distances of 2.979(2) Å and 2.927(2) Å, respectively. The second base pair is formed by two hydrogen bonds involving the protonated ring nitrogens of the 5-fluorocytosine rings and the carbonyl oxygens on neighbouring 5-fluorocytosine molecules (Figure 4.44). These are N-H...O hydrogen bonds with N...O distances of 2.779(2) Å and 2.826(2) Å, respectively and are therefore both of moderate strength. These base pairs alternate to form a stepped hydrogen bonded chain of alternating type A and B hydrogen bonding units (Figure 4.45, left). This chain is similar to that found in many of the cytosine molecular complexes where there is no hydrogen transfer.

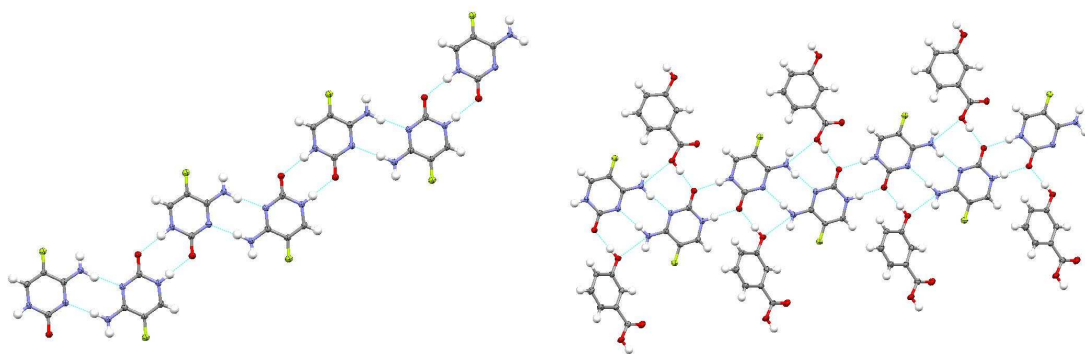


Figure 4.45. Left, the base-paired 5-fluorocytosine chain; right, 3-hydroxybenzoic acid molecules tied to each side of the chain in the molecular complex of 5-fluorocytosine and 3-hydroxybenzoic acid.

The chains have 3-hydroxybenzoic acid molecules tethered to each side of the chain. On one side of the chain, the benzoic acid molecule participates in two hydrogen bonds to the chain through the hydroxyl group, one as a hydrogen bond donor and one as a hydrogen bond acceptor (Figure 4.45, right). The first of these is an O-H...O hydrogen bond from the hydroxyl group to the carbonyl oxygen of a cytosine molecule in the chain. This has an O...O distance of 2.641(2) Å. The second is an N-H...O hydrogen bond from the amine group of a neighbouring cytosine molecule to the hydroxyl oxygen atom and has an N...O distance of 2.768(3) Å. On the other side of the chain, the carboxyl group of the 3-hydroxybenzoic acid molecule interacts with the chain through both oxygen atoms forming a moderate and a weak hydrogen bond (Figure 4.45, right). The moderate hydrogen bond is of O-H...O type from the protonated oxygen of the carboxyl group to the carbonyl oxygen of a cytosine molecule with an O...O distance of 2.678(2) Å. The weak C-H...O hydrogen bond is between an aromatic carbon atom on a neighbouring cytosine to the unprotonated oxygen on the carboxyl group. This hydrogen bond has a C...O distance of 3.025(4) Å. Each cytosine chain is planar, but neighbouring chains lie at approximately 35° to one another (Figure 4.46, left). Gaps are left between these strongly hydrogen bonded units.

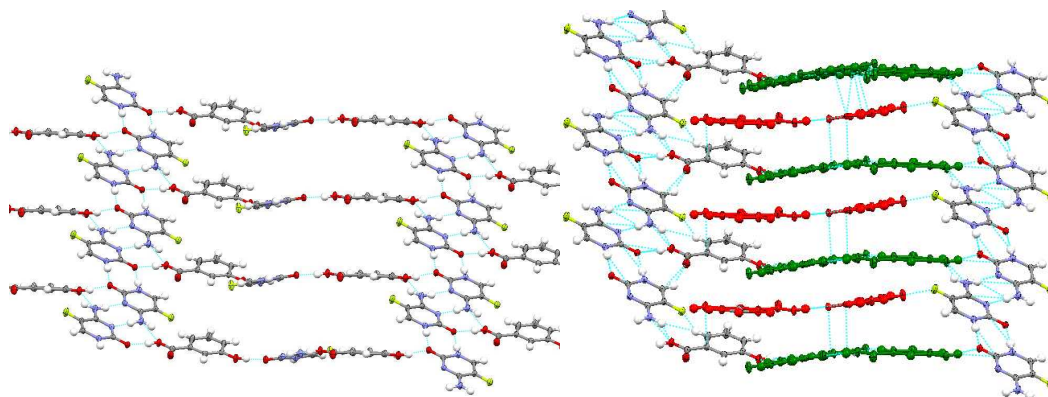


Figure 4.46. Left, change in orientation of adjacent chains and right, fluorine interactions in the 5-fluorocytosine and 3-hydroxybenzoic acid molecular complex. Green represents the chain held in place by hydrogen bonding and red represents the chain held in place via fluorine.

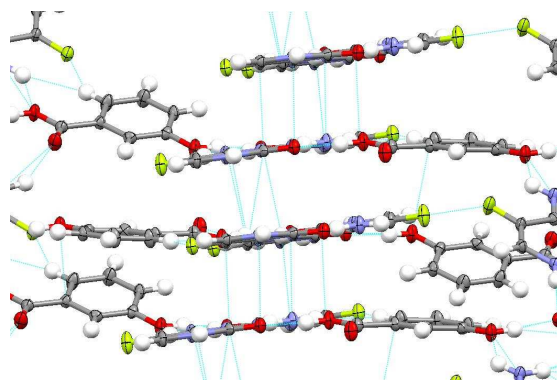


Figure 4.47. Stacking interactions in the 5-fluorocytosine and 3-hydroxybenzoic acid molecular complex.

The fluorine substituent on the 5-fluorocytosine molecules thus does not interrupt the base pairing, with the base-paired chains forming in the same manner as the non-fluorinated equivalent structures. However, the fluorine does affect the weaker interactions. Fluorine-fluorine interactions are found between cytosine molecules in different chains (Figure 4.46, right). The F...F distance is 2.801(2) Å which is shorter than the sum of the van der Waals radii of 2.94 Å. This fluorine interaction helps to link two equivalent hydrogen bonded networks together with the linked structure filling the gaps shown in Figure 4.46. Once these chains are housed within the “gaps” it is clear to see that once again there are base-stacking interactions between units, this time with 5-fluorocytosine molecules base stacked upon other 5-fluorocytosine molecules (Figure 4.48). These base stacking interactions are staggered relative to one another, with the more staggered interaction having the closest layer spacing (~ 3.089 Å) and the less staggered interaction having a layer spacing of approximately 3.213 Å. Interestingly, the closest stacking results in the cytosine molecules being stacked above the type B hydrogen bonded ring (Figure 4.48).

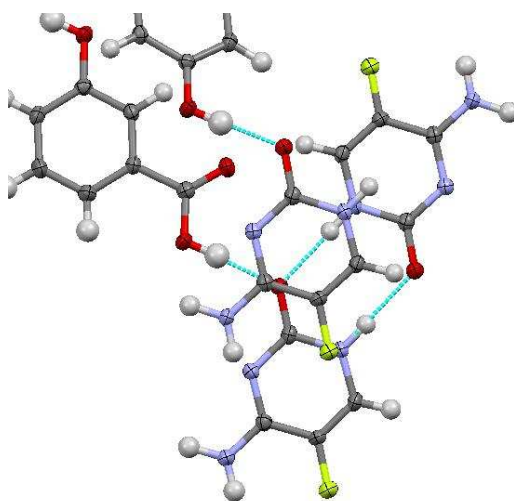


Figure 4.48 Staggered Stacking interactions corresponding to the closest interaction in the 5-fluorocytosine and 3-hydroxybenzoic acid molecular complex.

The effect of all of these hydrogen bonding motifs and fluorine interactions can be demonstrated by the view along the c-axis. This clearly shows the criss-cross pattern of these units, together with the base stacking (Figure 4.49).

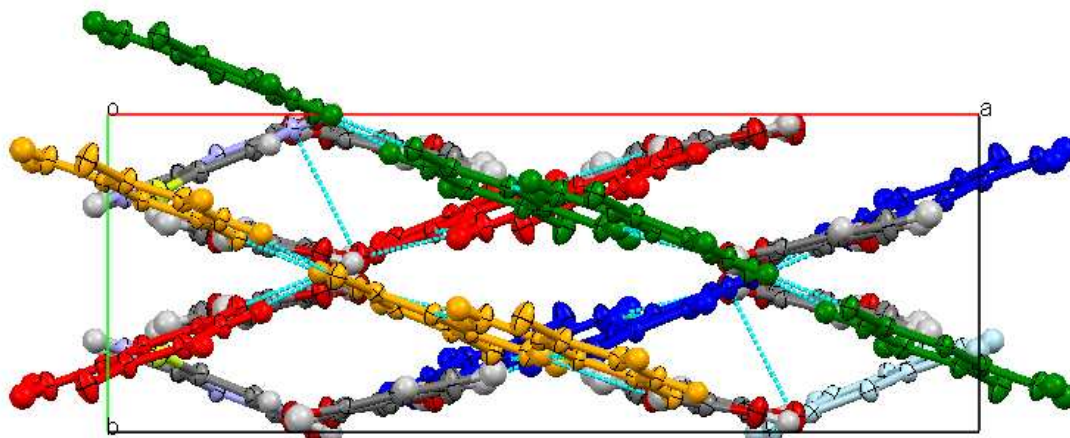


Figure 4.49. Criss-cross pattern formed in the 5-fluorocytosine and 3-hydroxybenzoic acid molecular complex.

4.8. 2:3:2 Molecular Complex of Cytosine and 4-Hydroxybenzoic Acid Hydrate

Cytosine and 4-hydroxybenzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.4mm x 0.4mm x 0.4mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 4.9.

The molecular complex forms in a 2:3:2 ratio of cytosine to 4-hydroxybenzoic acid to water. This is the only molecular complex obtained in this work where the ratio of cytosine to the benzoic acid is 2:3. There are seven molecules in the asymmetric unit, and the crystal belongs to the space group P-1. Whilst the ratio in the product is presented as 2:3:2, there is only partial proton transfer and so this leads to the creation of five different molecular species, two of which carry a charge; two neutral and one deprotonated 4-hydroxybenzoic acid molecules, a protonated cytosine (cytosinium) and a neutral cytosine, and two water molecules (Figure 4.50, left). The partial proton transfer allows the formation of a pseudo-Watson-Crick base paired dimer between the neutral and protonated cytosine molecules with three hydrogen bonds forming the supramolecular unit, one of which is charge assisted (Figure 4.50, right). This grouping is not possible for neutral cytosine molecules.

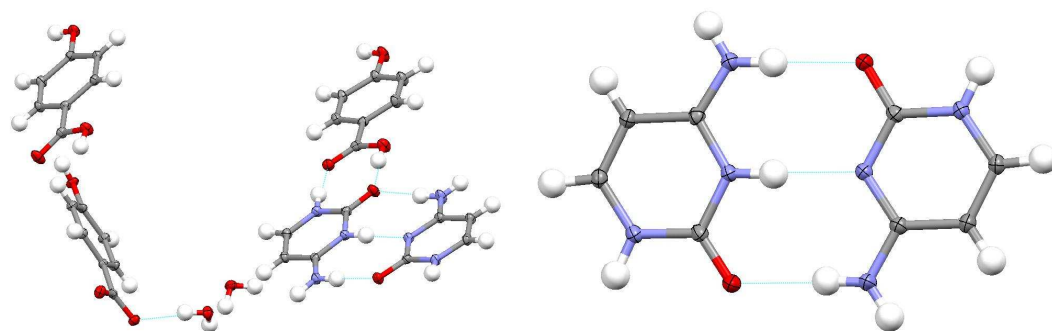


Figure 4.50. Left, the asymmetric unit showing proton transfer from one 4-hydroxybenzoic acid molecule to one cytosine molecule and right, the pseudo-Watson-Crick base pair found in the 2:3:2 molecular complex of cytosine 4-hydroxybenzoic acid hydrate.

The hydrogen bonds in the cytosine base pair are all of moderate strength. However, the three hydrogen bonds in the pseudo-Watson-Crick unit (Table 4.10) are shorter than those found in the molecular complex of ethyl-guanine and cytosine (EGMCYT10) [131]. However, the bonding found here is consistent with the structures discussed in Chapter 3 and elsewhere in this chapter. This is a consequence of the central hydrogen bond being charge assisted.

Table 4.10. Hydrogen bond distances for the pseudo-Watson-Crick bonding motif in the cytosine 4-hydroxybenzoic acid hydrate molecular complex.

Hydrogen Bond	D...A Distance
N-H---O	2.909(2) Å
N---H-N	2.848(2) Å
O---H-N	2.802(2) Å

In the structures described in Chapter 3, these pseudo-Watson-Crick base paired units form into hydrogen bonded chains of cytosine. In the molecular complexes of cytosine and 5-fluorocytosine with 5-fluorosalicylic acid hydrates, the cytosine chains are terminated by 5-fluorosalicylic acid molecules. However, in this case, hydrogen bonded chains are formed which include 4-hydroxybenzoic acid molecules, spacing the pseudo-Watson-Crick dimers (Figure 4.51). There are six hydrogen bonds formed from the base pair connecting four neutral 4-hydroxybenzoic acid molecules to a single base pair (Table 4.11). These benzoic acid molecules are arranged top to tail relative to one another. Two independent hydrogen bonded rings are formed between the carboxyl group of the benzoic acid and the carbonyl and ring nitrogen of both the cytosine and cytosinium molecules (A,B,C,D, Figure 4.51). The other two hydrogen bonds tie two separate 4-hydroxybenzoic acids to the base pair (E and F, Figure 4.51). These hydrogen bonds involve the oxygen of the hydroxyl

group on the benzoic acid and one of the hydrogen atoms on the amine group of the cytosine.

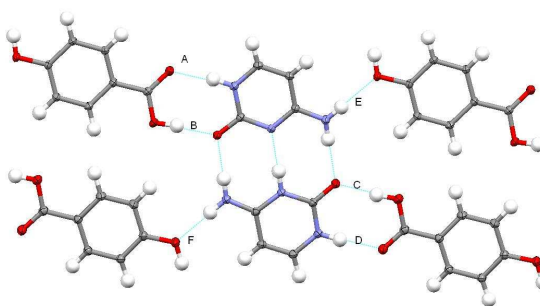


Figure 4.51. The hydrogen bonded chain formed in the 4-hydroxybenzoic acid and cytosine hydrate molecular structure.

Table 4.11. Hydrogen bonds involving the pseudo-Watson-Crick base pair in 4-hydroxybenzoic acid and cytosine hydrate molecular structure.

Hydrogen Bond	D...A Distance
A (N-H...O, ring N to carboxyl)	2.782(3) Å
B (O-H...N, carboxyl to carbonyl)	2.671(3) Å
C (O-H...O carboxyl to ketone)	2.669(4) Å
D (N-H...O ring N to carboxyl)	2.755(2) Å
E (N-H...O amine to hydroxyl)	2.929(1) Å
F(N-H...O amine to hydroxyl)	2.817(3) Å

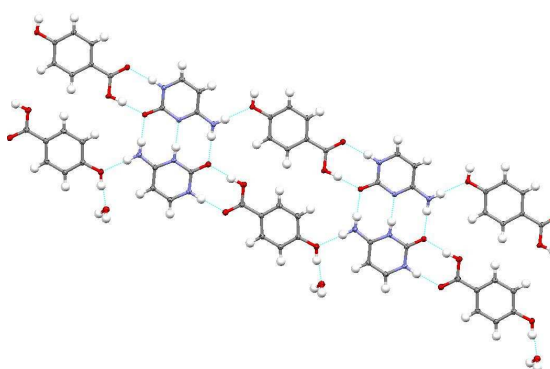


Figure 4.52. Extended chain with water molecules tied to one side of the chain in the cytosine 4-hydroxybenzoic acid hydrate molecular complex.

This chain propagates along the *ab* diagonal (Figure 4.52). The hydroxyl hydrogen atom points away from the chain and is able to act as a hydrogen bond donor. On one side of the chain a hydrogen bond is formed to a water molecule and is of moderate strength with an O...O distance of 2.579(2) Å.

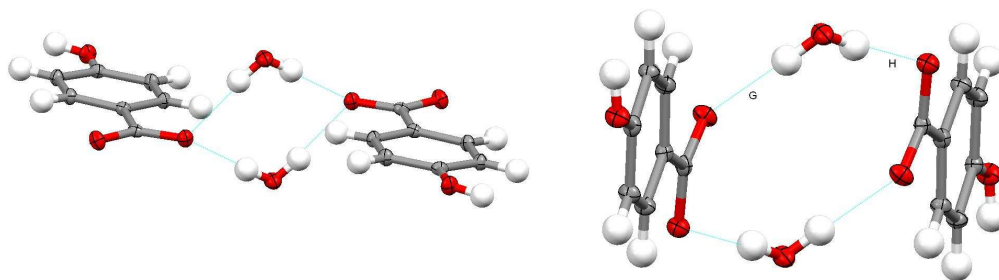


Figure 4.53. The two four membered hydrogen bonded rings of two water molecules and two 4-hydroxybenzoate molecules in the cytosine 4-hydroxybenzoic acid hydrate molecular complex.

This water molecule hydrogen bonds to two 4-hydroxybenzoate molecules, acting as a hydrogen bond donor with O...O distances of 2.792(2) Å and 2.769(2) Å, respectively. An $R_4^2(8)$ hydrogen bonded ring motif is formed from two water molecules and two 4-hydroxybenzoate molecules (Figure 4.53, left). This unit then acts as a spacer between chains which lie in approximately the same plane. The carboxylate groups in this ring connect layers of the chains through a hydrogen bond to the hydroxyl group on the other side of the chain. This hydrogen bond is of moderate strength and has an O...O distance of 2.614(2) Å.

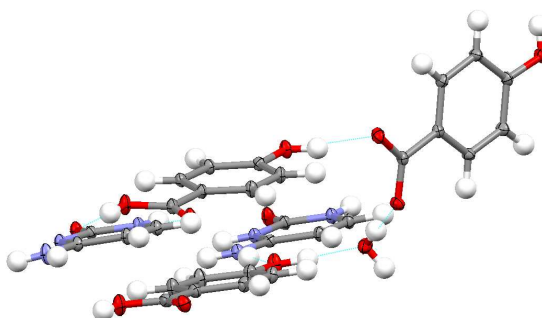


Figure 4.54. Charged 3-hydroxybenzoic acid molecule linking the layers in the cytosine 4-hydroxybenzoic acid hydrate molecular complex.

The hydroxyl group on the 4-hydroxybenzoate molecule forms a hydrogen bond to a second water molecule with an O...O distance of 2.834(2) Å. This water molecule acts as a hydrogen bond donor to two carboxylate groups and creates hydrogen bonded rings, $R_4^4(12)$, consisting of two water molecules and two benzoic acid molecules (Figure 4.52, right). Both hydrogen bonds are of moderate strength and have O...O distances of 2.681(2) Å (G in Figure 4.53) and 2.849(2) Å (H in Figure 4.53), respectively. These rings are interconnected and form a chain constructed of deprotonated benzoic acid molecules and water molecules, which propagates along the *a*-axis in an alternating fashion (Figure 4.55).

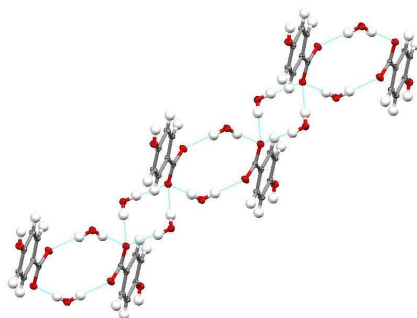


Figure 4.55. Combination of two four membered rings forming an extended chain in the cytosine and 4-hydroxybenzoic acid hydrate molecular complex.

There is also base stacking, between pairs of cytosine and pairs of 4-hydroxybenzoic acid molecules. This structure contains some base stacking that has pseudo-Watson-Crick base pairs stacked on top of each other with charged cytosine molecules on top of the neutral cytosine and the vice versa on the other side. This stacking has a distance of 3.063 Å, which is in line with the previously discussed structures that contain the pseudo-Watson-Crick bonding motif.

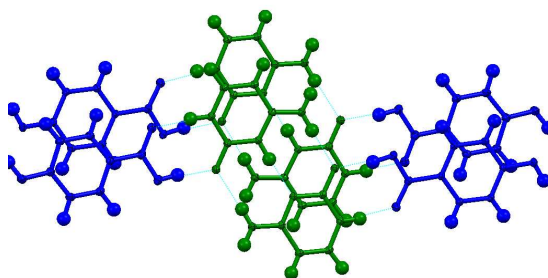


Figure 4.56. Stacking interactions in the cytosine and 4-hydroxybenzoic acid hydrate molecular complex.

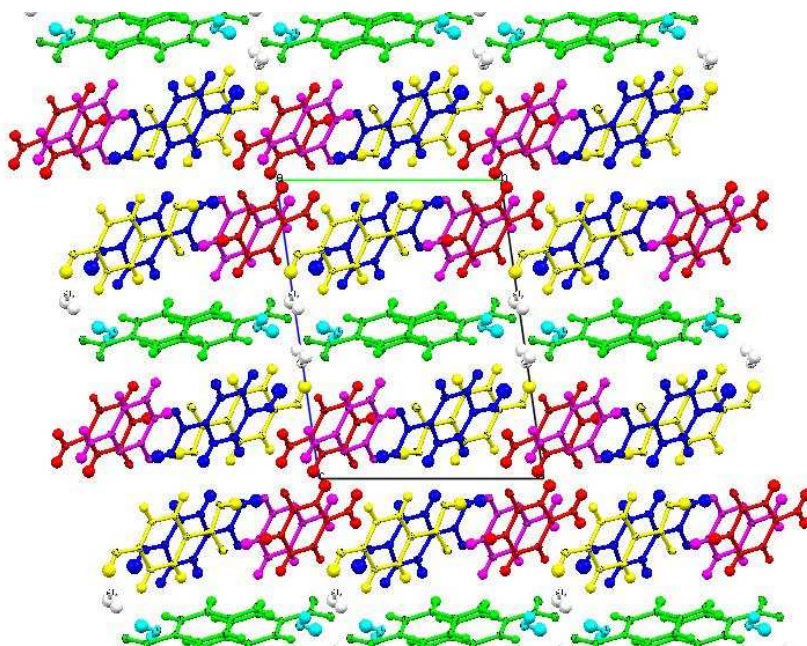


Figure 4.57. Packing in the cytosine and 4-hydroxybenzoic acid hydrate molecular complex.

There is one more stacking interaction between two 4-hydroxybenzoic acid molecules (Figure 4.56). The stacking interaction between the regions of the 4-hydroxybenzoic acid

molecules is in the range of approximately 3.441–3.781 Å. Figure 4.57 shows the 4-hydroxybenzoic acid and the water molecules working to link separate chains together with the chains of cytosine and 4-hydroxybenzoic acid molecules stacked on top of each other in a staggered fashion. In the stacking area the neutral cytosine molecule once again stacks on top of a cytosinium molecule.

4.9. 1:1 Molecular Complex of 5-Fluorocytosine and 4-Hydroxybenzoic Acid

5-Fluorocytosine and 4-hydroxybenzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.2mm x 0.3mm x 0.4mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 4.9.

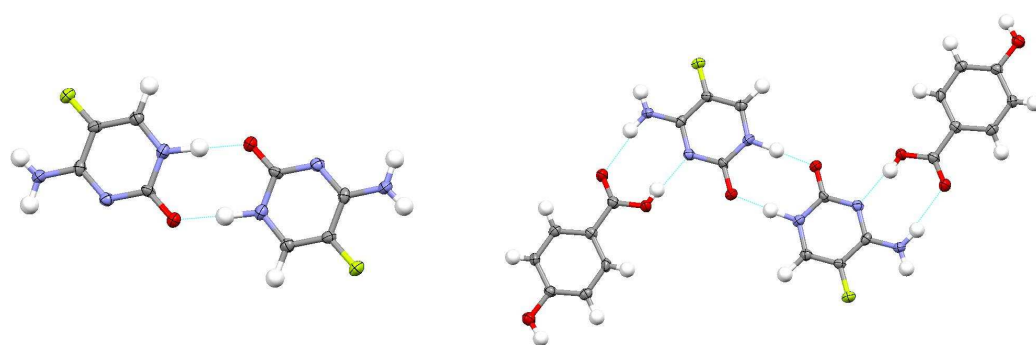


Figure 4.58. Left, type B hydrogen bonding motif and right, four molecule basic building block in the 5-fluorocytosine and 4-hydroxybenzoic acid molecular complex.

This structure contains no hydrogen transfer and crystallises in a 1:1 ratio of 5-fluorocytosine to 4-hydroxybenzoic acid. The 5-fluorocytosine molecules are therefore capable of forming both the type A and type B pseudo Hoogsten base pair made up of two hydrogen bonds. The different crystallisation ratio from that found in the 3-hydroxybenzoic acid 5-fluorocytosine molecular complex results in a different construction of the primary hydrogen bonded motifs. Here, only one of these types of base pair are formed between 5-fluorocytosine molecules, type B (Figure 4.58). The base pair is made up two equivalent N-H...O hydrogen bonds between the protonated ring nitrogens and the carbonyl oxygens on neighbouring cytosine molecules. These hydrogen bonds are symmetry equivalent with an inversion centre located in the centre of the hydrogen bonded ring, and are therefore both of moderate strength with an N...O distance of 2.798(2) Å. An extended chain of cytosine molecules does not occur; 4-hydroxybenzoic acid molecules form hydrogen-bonded rings to the other sites at either end of the cytosine base pair.

Interestingly, the hydroxyl groups do not act as hydrogen bond donors to any other molecules and so the chain is terminated. This creates a distinct four molecule supramolecular unit (Figure 4.58, right).

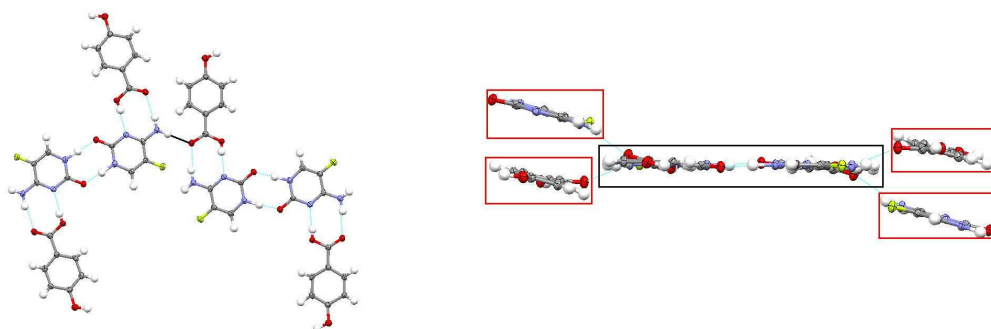


Figure 4.59. Left, two separate building blocks linked via a single hydrogen bond (black lines) and right, Orientation of the molecules hydrogen bonded to four molecule building block via single hydrogen bonds in the 5-fluorocytosine and 4-hydroxybenzoic acid molecular complex.

The hydrogen-bonded ring between the 5-fluorocytosine and 4-hydroxybenzoic acid molecules is made up of two hydrogen bonds: the first is an O-H \cdots N hydrogen bond from the protonated oxygen of the carboxyl group to the unprotonated ring nitrogen of the cytosine; the second is an N-H \cdots O hydrogen bond from the amine group of the cytosine to the unprotonated oxygen on the carboxyl group. The two hydrogen bonds are of moderate strength and have D \cdots A lengths of 2.656(2) Å and 2.973(2) Å, respectively. The carboxyl group of the 4-hydroxybenzoic acid molecule is not planar with respect to the benzene ring plane; it is twisted out of the plane by approximately 15°. The amine group of the cytosine is involved in a second hydrogen bond to the unprotonated oxygen of the carboxyl group. The twist of the carboxyl group allows this hydrogen bond to be shorter and is thus likely to be the driving force behind the twisting. This hydrogen bond is of moderate strength and has an N \cdots O distance of 2.983(2) Å. This hydrogen bond acts as a link between different supramolecular units of the 4-molecule building block (Figure 4.59, left). From each hydrogen-bonded ring there is a hydrogen bond that goes above the pair and one that goes below the original pair (Figure 4.59, right).

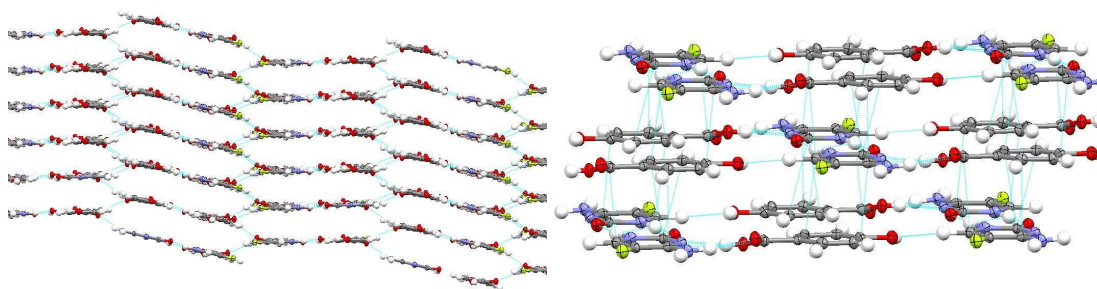


Figure 4.60. Left, layered nature of structure and right, stacking interactions in structure of 5-fluorocytosine and 4-hydroxybenzoic acid molecular complex.

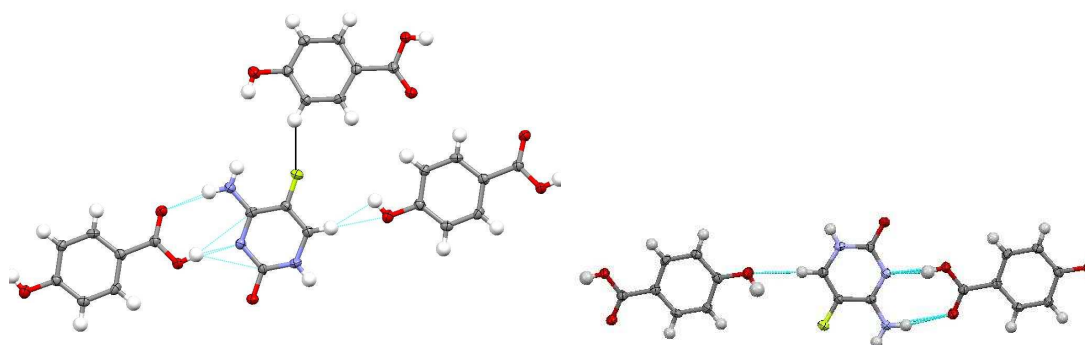


Figure 4.61. Left, weak C-H...F hydrogen bond and right, weak C-H...O hydrogen bond in the 5-fluorocytosine and 4-hydroxybenzoic acid molecular complex.

The overall structure takes a layered form (Figure 4.60), in which base stacking plays a significant role. The fluorine takes part in one weak hydrogen bond and this occurs between the layers (Figure 4.61, left). This weak hydrogen bond comes from an aromatic hydrogen atom of the 4-hydroxybenzoic acid molecule and is of C...F length of 3.497(2) Å. A weak hydrogen bond also exists between the oxygen atom of the hydroxyl group of the 4-hydroxybenzoic acid and an aromatic hydrogen atom of the 5-fluorocytosine (Figure 4.61, right). This hydrogen bond has a C...O length of 3.232(2) Å. It is unusual to see the hydroxyl group not acting as a proton donor, but not unusual for it to act as a hydrogen bond acceptor.

There is stacking of alternating 5-fluorocytosine and 4-hydroxybenzoic acid molecules (Figure 4.62). Here the stacking interactions are of approximate distance of 3.249 Å from the cytosine to a hydroxybenzoic acid above and 3.305 Å from a hydroxybenzoic acid to the cytosine above it.

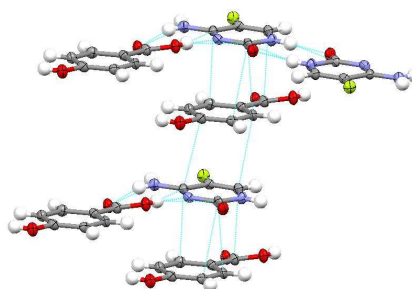


Figure 4.62. Stacking interactions in the 5-fluorocytosine and 4-hydroxybenzoic acid molecular complex. Three stacking interactions show with two being equivalent.

4.10. Structural Comparison of Molecular Complexes Formed Between Cytosine and 5-Fluorocytosine with Mono-Substituted Hydroxybenzoic Acids

The predominant structural trends obtained in the molecular complexes of the mono-substituted hydroxybenzoic acid derivatives with cytosine and 5-fluorocytosine are dominated by the proton transfer in the structures. The structures belong to three well-defined categories depending on the level of proton transfer in the structure: neutral, mixed charge, and fully charged molecular complexes. The fluorine interactions play a significant role in the assembly of these units in combination with base-stacking interactions.

4.10.1. The Effect of Proton Transfer

ΔpK_a values can be a helpful guide when predicting whether proton transfer will occur. The general rule for the ΔpK_a values between the components is that if this difference is greater than 3 there should be hydrogen transfer, the area of possible hydrogen transfer is between 0 and 3 and the region of no hydrogen transfer for $\Delta pK_a \leq 0$ [135].

When considering the ΔpK_a values for the series of molecular complexes presented in this chapter, it becomes clear that the level of proton transfer can be explained reasonably well by this liquid-phase property. The structures that would be predicted to be most likely to have hydrogen transfer are the structures involving salicylic acid (pK_a 2.98) and 5-fluorosalicylic acid as these are the molecules that have the biggest difference compared to those of 5-fluorocytosine (pK_a 3.26) and cytosine (pK_a 4.60). The structures in this series follow this trend as all of the structures that have salicylic acid as one of the co-molecules, have at least partial transfer of a hydrogen atom, with three out of the five structures having total transfer with respect to the cytosine molecule (Table 4.12).

Table 4.12 The ΔpK_a and the % transfer of a proton with respect to the cytosine molecules

Molecular complex	ΔpK_a	Hydrogen Transfer with respect to cytosine molecules
Cytosinium Salicylate	1.62	100%
Salicylic acid and 5-Fluorocytosine	0.28	100%
5-Fluorosalicylic acid and 5-Fluorocytosine acid Hydrate	0.58	50%
5-Fluorosalicylic acid and Cytosine	1.92	100%
5-Fluorosalicylic acid and Cytosine Hydrate	1.92	50%
3-hydroxybenzoic acid and Cytosine	0.52	100%
3-hydroxybenzoic acid and 5-Fluorocytosine	-0.82	0%
4-hydroxybenzoic acid and Cytosine	0.03	50%
4-hydroxybenzoic acid and 5-Fluorocytosine	-1.31	0%

The 3-hydroxybenzoic acid molecular complexes are the next most likely to be involved in hydrogen transfer. The only molecular complex that contains hydrogen transfer is that of 3-hydroxybenzoic acid and cytosine; the ΔpK_a for this is 0.52 putting it in the region for possible hydrogen transfer; this structure does, in fact, exhibit total hydrogen transfer with respect to the cytosine. The pK_a value of 5-fluorocytosine is lower than that of the 3-hydroxybenzoic acid which means that this structure is not expected to contain any hydrogen transfer and the molecular complex obtained is neutral.

The pK_a value for 4-hydroxybenzoic acid is high (4.57) and thus hydrogen transfer is less likely. Here there is partial hydrogen transfer in the cytosine and 4-hydroxybenzoic acid structure, with the ΔpK_a value very close to 0 but still positive. The molecular complex with 5-fluorocytosine does not contain any hydrogen transfer as would be expected for a negative ΔpK_a value.

4.10.2. Complexes Containing Neutral Molecules Only

In this series of nine structures, there are only two structures that do not contain hydrogen transfer: 3-hydroxybenzoic acid and 4-hydroxybenzoic acid with 5-fluorocytosine. These structures all show 5-fluorocytosine molecules linked together by type A and/or B base pair motifs in the same way as the neutral complexes with fluorosubstituted benzoic acids and benzoic acid itself (Chapter 3).

4.10.2.1. Comparison Between the Molecular Complexes of 5-Fluorocytosine with Benzoic Acid and 3-Hydroxybenzoic Acid

Despite the different crystallisation ratios between the 5-fluorocytosine benzoic acid 1:1, Section 3.6.1) and 5-fluorocytosine 3-hydroxybenzoic acid (2:1) molecular complexes, there are some structural similarities between the two. Both show extended chains of 5-fluorocytosine molecules constructed through alternating type A and B hydrogen bonding motifs (Figure 4.63). Neutral benzoic acid or 3-hydroxybenzoic acid molecules are tethered to each side of this chain (Figure 4.64). Since the benzoic acid molecule only has one hydrogen bond donor site, the benzoic acid molecules are tethered to the cytosine chain in the same way on both sides (Figure 4.64, left). The carboxylic acid group of the benzoic acid molecule acts as a hydrogen bond donor in a moderate strength hydrogen bond ($O\cdots O$ distance of 2.671(2) Å) and as a hydrogen bond acceptor in a $C-H\cdots O$ hydrogen bond ($C\cdots O$ distance of 3.064(2) Å). 3-hydroxybenzoic acid however, has a second hydrogen bond donor site, the hydroxyl group, and thus the coordination to the 5-

fluorocytosine chain is only through the carboxyl group on one side of the chain and on the other side, the coordination is through the hydroxyl group (Figure 4.64, right). The carboxyl group again forms a moderate and a weak hydrogen bond with $O\cdots O$ and $C\cdots O$ distances of 2.679(2) Å and 3.025(4) Å, respectively, similar to those of the cytosine benzoic acid molecular complex. The $O-H\cdots O$ hydrogen bond involving the hydroxyl group is of similar strength to the equivalent $O-H\cdots O$ hydrogen bond from the carboxyl group and has an $O\cdots O$ distance of 2.643(3) Å. This additional hydrogen bond donor site allows hydrogen bonded layers to be generated in the 5-fluorocytosine 3-hydroxybenzoic acid molecular complex which are not possible in the 5-fluorocytosine benzoic acid molecular complex.

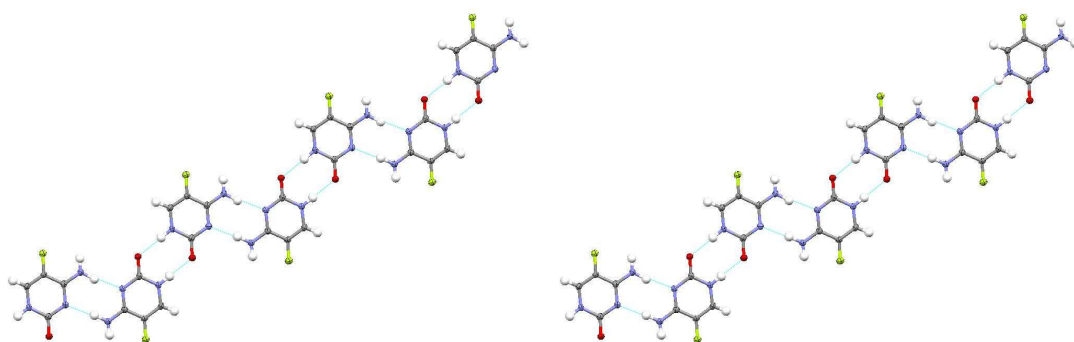


Figure 4.63. Chains of 5-fluorocytosine with alternating type A and B hydrogen bond motifs in the molecular complexes of 5-fluorocytosine with benzoic acid (left) and 3-hydroxybenzoic acid (right).

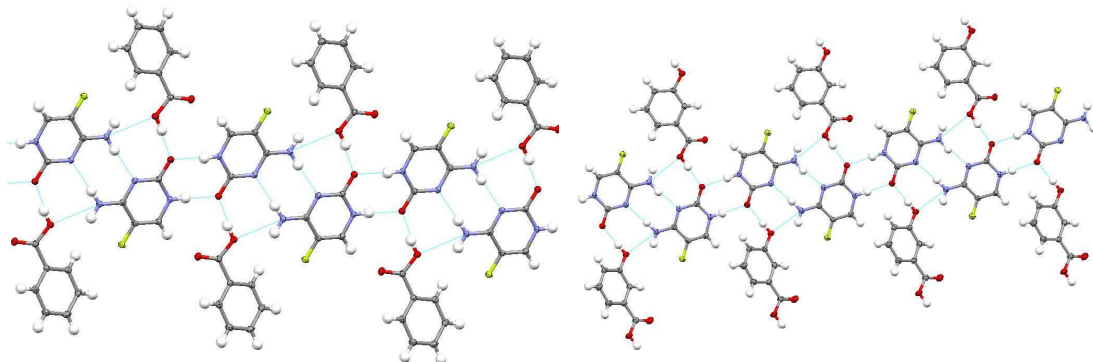


Figure 4.64. Benzoic acid (left) and 3-hydroxybenzoic acid (right) tied to the 5-fluorocytosine chains in their respective structures.

There are also some similarities in the weaker interactions. Both molecular complexes show $C-H\cdots F$ hydrogen bonds between an ortho aromatic hydrogen atom on the benzoic acid (or 3-hydroxybenzoic acid) molecule and the fluorine atom on the 5-fluorocytosine with $C\cdots F$ distances of 3.253(2) Å and 3.297(2) Å, respectively. In combination with an $N-H\cdots O$ moderate strength hydrogen bond between the amine group of the 5-fluorocytosine and the protonated oxygen atom of the carboxylic acid group on the benzoic acid ($N\cdots O$ distance 2.900(2) Å) or 3-hydroxybenzoic acid ($N\cdots O$ distance 2.888(2) Å)

molecule, these form hydrogen bonded rings within layers (Figure 4.65). These hydrogen bonds are of a similar strength.

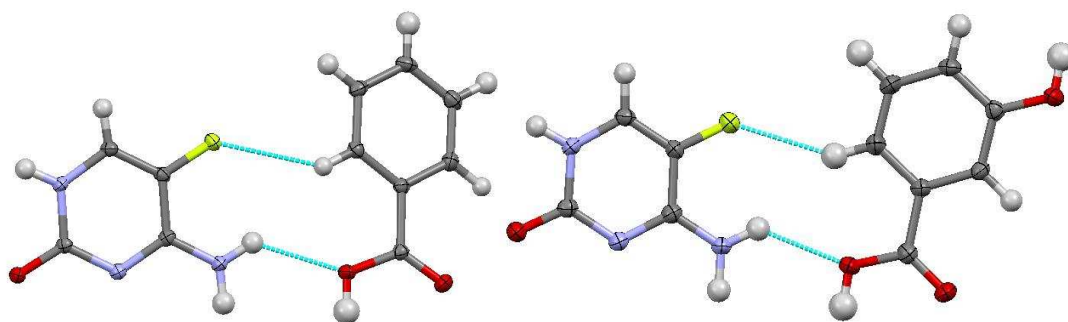


Figure 4.65. The similar hydrogen bonded units involving C-H...F hydrogen bonds in the left, 5-fluorocytosine benzoic acid and right, 5-fluorocytosine 3-hydroxybenzoic acid molecular complexes.

The different crystallisation ratio allows the molecular complex of 5-fluorocytosine and 3-hydroxybenzoic acid to form F...F interactions, whereas the molecular complex of 5-fluorocytosine with benzoic acid shows only C-H...F hydrogen bonds. Both also show stacking of cytosine chains on top of one another, staggered with respect to one another.

4.10.2.2. The Contrasting 4-Hydroxybenzoic Acid 5-Fluorocytosine Molecular Complex

The 4-hydroxybenzoic acid and 5-fluorocytosine structure is quite different from the molecular complex of 5-fluorocytosine and 3-hydroxybenzoic acid, crystallising in a 1:1 ratio. Two cytosine molecules hydrogen bond to one another through a type B pseudo-Hoogsten motif as found in the other neutral 5-fluorocytosine molecular complexes. However, instead of expanding through a type A hydrogen bond motif, the chain is terminated by formation of a hydrogen bonded ring with 4-hydroxybenzoic acid molecules (Figure 4.66). There are weak C-H...F hydrogen bonds between units and base stacking is still prevalent but in this case, 5-fluorocytosine molecules tend to stack above 4-hydroxybenzoic acid molecules and vice versa.

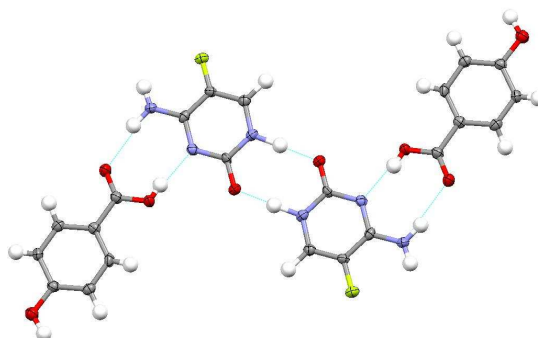


Figure 4.66. Termination of 5-fluorocytosine chain via formation of a hydrogen bonded ring in the 5-fluorocytosine and 4-hydroxybenzoic acid molecular complex.

4.10.3. Molecular Complexes Containing Charged Species Only

There are four structures in this series that contain full hydrogen transfer to the unprotonated nitrogen of the cytosine from the carboxyl group of the benzoic acid. This creates cytosinium and benzoate ions. These molecular complexes are: cytosinium salicylate, 5-fluorocytosinium salicylate, cytosinium 5-fluorosalicylate, and cytosinium 3-hydroxybenzoate hemihydrate. All of these structures form in a 1:1 ratio of cytosinium to the benzoate. The molecular complex cytosinium 3-hydroxybenzoate hemihydrate is the only hydrated complex and this forms in a 2:2:1 ratio of cytosinium to 3-hydroxybenzoate to water. All four structures adopt the same primary hydrogen bonding pattern of the heterodimer hydrogen bonded ring (Figure 4.67). This hydrogen bonding motif involves two hydrogen bonds, one from the amine group of the cytosine to one of the oxygen atoms of the carboxylate group of the benzoate and the other between the protonated nitrogen of the cytosinium and the other oxygen of the carboxylate group of the benzoate (Table 4.13).

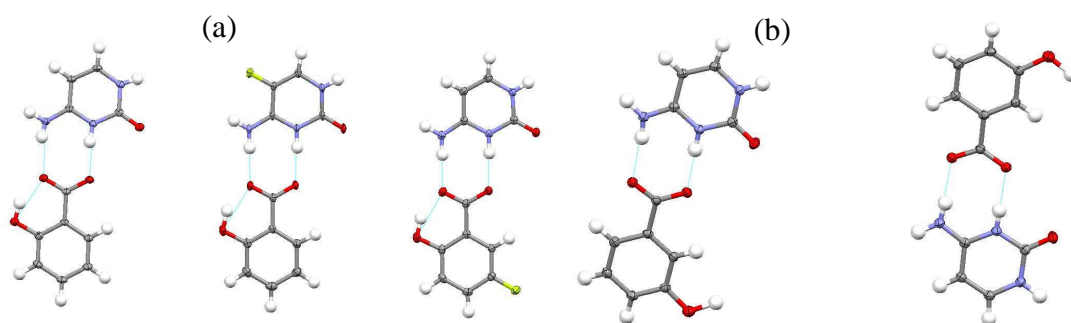


Figure 4.67. Primary hydrogen bonding motif in the four related charged structures; (a) left, cytosinium salicylate middle, 5-fluorocytosinium salicylate right, cytosinium 5-fluorosalicylate; (b) two independent heterodimers in cytosinium 3-hydroxybenzoate hemihydrate.

Table 4.13. Hydrogen bonds of the primary hydrogen bonding motif, the heterodimer, in the fully charged molecular complexes.

Molecular Complex	N-H---O (amine)	N-H---O (ring N)
Cytosinium Salicylate	2.748(2) Å	2.724(2) Å
5-Fluorocytosinium Salicylate	2.752(2) Å	2.770(2) Å
Cytosinium 5-Fluorosalicylate	2.759(2) Å	2.784(2) Å
Cytosinium 3-hydroxybenzoate dimer 1	2.773(2) Å	2.787(2) Å
Cytosinium 3-hydroxybenzoate dimer 2	2.773(2) Å	2.741(2) Å

The hydrogen bonding environment around the heterodimers is the same for all three anhydrous molecular complexes (Figure 4.68). There are three hydrogen bonds from the cytosinium and one hydrogen bond involving the salicylate; there are however, only two unique hydrogen bonds (Table 4.14). The hydrogen bonds in the two complexes involving salicylate are approximately the same; however, the hydrogen bonds in the complex with

5-fluorosaliclyate are considerably longer. The presence of the fluorine on the 5-fluorosaliclyate molecule may be playing a significant part in this.

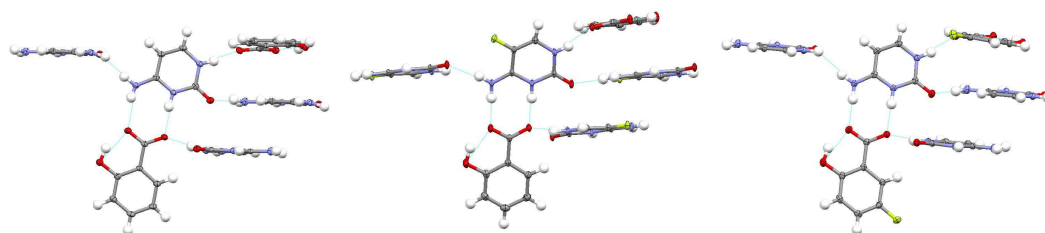


Figure 4.68. Hydrogen bonding environment of the cytosinium molecule in: left, cytosinium salicylate; middle, 5-fluorocytosinium salicylate; and right, cytosinium 5-fluorosaliclyate.

Table 4.14 Hydrogen bonds involving the primary hydrogen bonding motif in the three anhydrous fully charged molecular complexes.

Molecular Complex	N-H...O (ring to carboxyl) (Å)	N-H...O (amine to carbonyl) (Å)
Cytosinium salicylate	2.693(2)	2.872(2)
5-Fluorocytosinium salicylate	2.700(2)	2.871(2)
Cytosinium 5-fluorosaliclyate	2.722(2)	3.041(2)

All three of these molecular complexes form the same cytosinium chain constructed from adjacent heterodimers of cytosinium and salicylate aligned almost perpendicular to one another (Figure 4.69). The links to neighbouring chains are also the same in all three structures. This takes the form of a hydrogen bond from a protonated ring nitrogen of the cytosinium to one of the oxygen atoms of the carboxylate group (Figure 4.70).

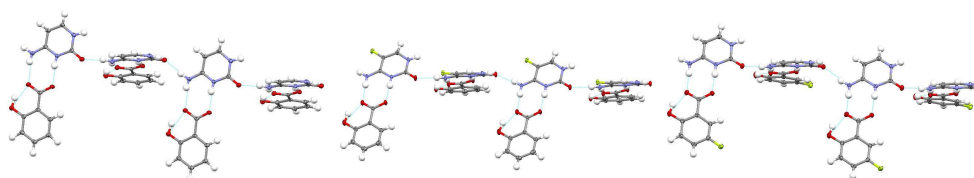


Figure 4.69. Hydrogen bonded chains of heterodimers in: left, cytosinium salicylate; middle; cytosinium 5-fluorosaliclyate; right, 5-fluorocytosinium salicylate.

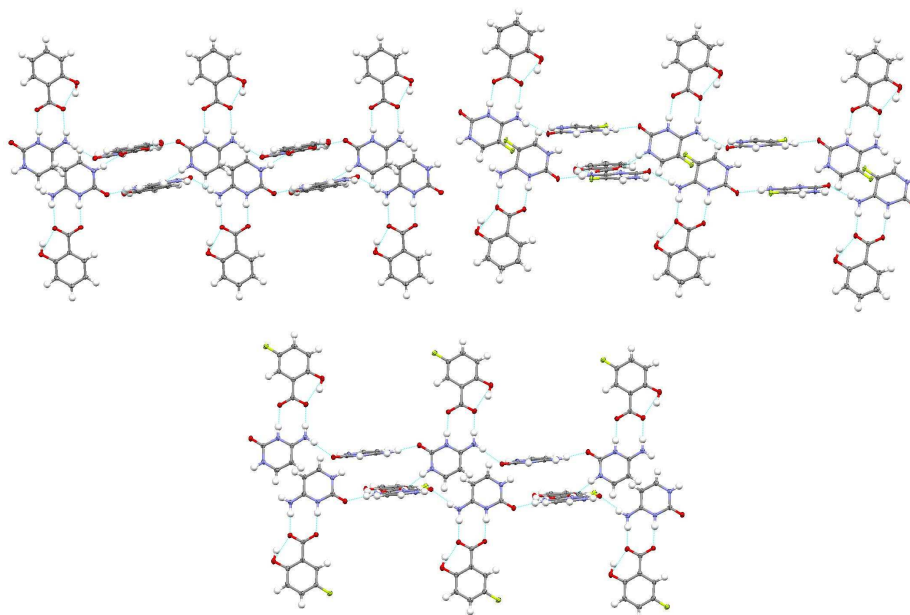


Figure 4.70. The linking of the cytosinium chains in the molecular complexes of: top, left cytosinium salicylate; top right, cytosinium 5-fluorosaliclyate; bottom, 5-fluorocytosinium salicylate.

The main difference between the structures is the positioning of the fluorine on the molecular components and the interactions formed as a consequence of this. This explains why there are differences in the unit cell parameters (Table 4.15).

Table 4.15. Unit cells of the three anhydrous fully charged molecular complexes.

Crystal Structure	Cytosinium salicylate	5-Fluorocytosinium salicylate	Cytosinium 5-fluorosaliclyate
Space Group	P2 ₁ /n	P2 ₁ /n	P2 ₁ /n
<i>a</i> (Å)	9.6425(3)	7.8518(17)	8.3534(7)
<i>b</i> (Å)	8.4698(3)	10.6457(24)	10.1434(7)
<i>c</i> (Å)	13.9021(4)	13.3065(29)	13.1203(10)
α (°)	90	90	90
β (°)	109.887(2)	100.041(15)	101.860(4)
γ (°)	90	90	90

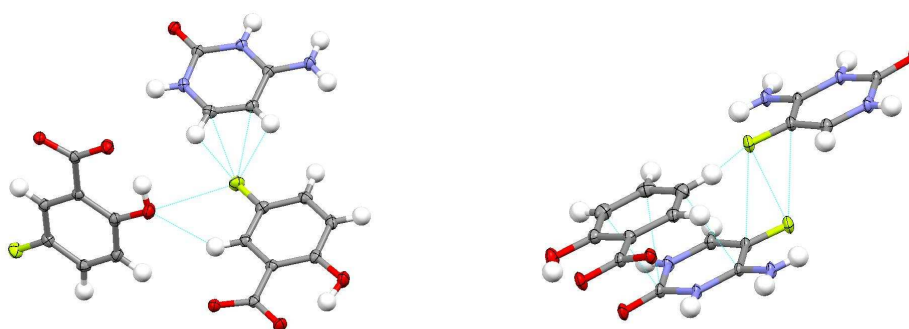


Figure 4.71. Fluorine interactions in left, cytosinium 5-fluorosaliclyate and right 5-fluorocytosinium salicylate molecular complexes.

There are two fluorine interactions in the cytosinium 5-fluorosaliclyate molecular complex, one of which involves the cytosinium molecule and one between two salicylate molecules (Figure 4.71, left). The interaction with the cytosinium can be classed as a bifurcated DDHHA C-H...F weak hydrogen bond with C...F distances of 3.027(2) Å and 3.079(2) Å.

The cytosinium molecule lies in the same plane as the salicylate. The other fluorine interaction is between the fluorine and the hydroxyl oxygen of a nearby salicylate molecule. This O...F distance is 2.925(2) Å and is somewhat shorter than the sum of the van der Waals radii of oxygen and fluorine which is 2.99 Å. This pulls the molecules closer together and contracts the unit cell in two of the three axes in comparison to the non-fluorinated cytosinium salicylate molecular complex. The 5-fluorocytosinium salicylate molecular complex also has an F...F interaction (Figure 4.71, right). This is between two 5-fluorocytosinium molecules that are in two separate planes parallel to one another. This interaction has an F...F distance of 2.744(2) Å (c.f. sum of the van der Waals radii of 2.94). There is a weak C-H...F hydrogen bond with a C...F distance of 3.278(2) Å. This is weaker than that found in the cytosinium 5-fluorosaliclate molecular complex.

The cytosinium 3-hydroxybenzoate molecular complex does not fit this pattern as the structure contains water molecules, and also because the hydroxyl group is available to form intermolecular hydrogen bonds. There are two independent heterodimers and these have different hydrogen bond environments. Five hydrogen bonds are still formed per cytosinium molecule, two of which comprise the heterodimer hydrogen bonded ring. However, instead of forming two hydrogen bonds to cytosinium molecules and one to a benzoate, here there is only one hydrogen bond to another cytosinium molecule. The additional two hydrogen bonds are formed to two 3-hydroxybenzoate molecules in one case, and one 3-hydroxybenzoate and one water in the other. In common with the other structures, there are still hydrogen bonds between a protonated ring nitrogen of the cytosine to a nearby 3-hydroxybenzoic acid molecule with N...O distances of 2.688(2) Å (dimer 1) and 2.760(2) Å (dimer 2), shorter than those found in the other molecular complexes. There is also a hydrogen bond from a protonated nitrogen to a carboxyl oxygen of another cytosine. This bond is of moderate strength and is of length 2.760(2) Å (dimer 1) and 2.718(2) Å (dimer 2).

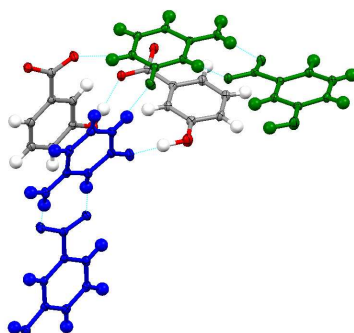


Figure 4.72. The hydrogen bonds from the heterodimer (dimer 1 green and dimer 2 blue) in the cytosinium 3-hydroxybenzoate hemihydrate molecular complex

When there is 100% transfer of a hydrogen atom creating solely charged species in the molecular complex, the predominant hydrogen bonding motif is the heterodimer hydrogen bonded ring between a cytosine and the carboxylate group of the substituted benzoic acid. None of the fluorine substitution, the hydroxyl group or the introduction of a water molecule interrupts this basic hydrogen bonding motif.

4.10.4. Molecular Complexes Showing Mixed Charge Molecular Components

In this case, both neutral and protonated cytosine molecules are found within the molecular complex and these molecular complexes are all hydrates. These are: cytosine and 4-hydroxybenzoic acid hydrate, 5-fluorocytosine and 5-fluorosalicylic acid hydrate, cytosine and 5-fluorosalicylic acid hydrate. When there is mixed hydrogen transfer, the predominant hydrogen bonding motif is the pseudo-Watson-Crick base pair. This hydrogen bonding motif is only possible if both cytosine and cytosinium molecules co-exist in the molecular complex (Figure 4.73). The hydrogen bond strengths are similar in all of these molecular complexes (Table 4.16).

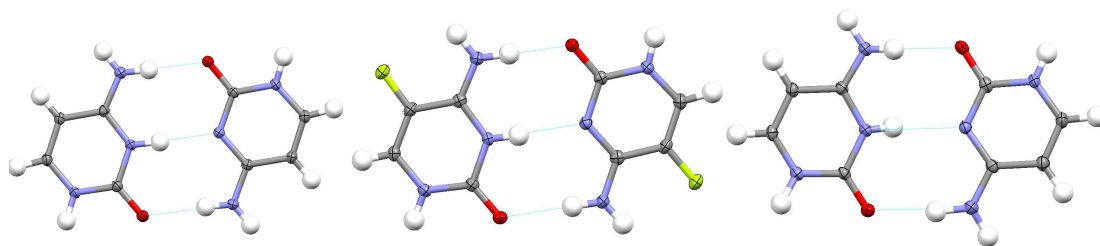


Figure 4.73. Pseudo-Watson-Crick base pair in the, left cytosine 4-hydroxybenzoic acid hydrate molecular complex, middle 5-fluorocytosine 5-fluorosalicylic acid hydrate molecular complex and right, cytosine 5-fluorosalicylic acid hydrate molecular complex

Table 4.16. Hydrogen bonds of the pseudo-Watson-Crick hydrogen bonding motif

	1. N-H---O (Å)	2. N---H-N (Å)	3. O---H-N (Å)
Cytosine and Benzoic acid	2.771(3)	2.844(2)	2.914(2)
Cytosine and 4-hydroxybenzoic acid hydrate	2.805(2)	2.848(3)	2.909(2)
5-fluorocytosine and 5-fluorosalicylic acid hydrate	2.740(2)	2.837(3)	2.967(2)
Cytosine and 5-fluorosalicylic acid hydrate	2.767(2)	2.852(2)	2.943(3)

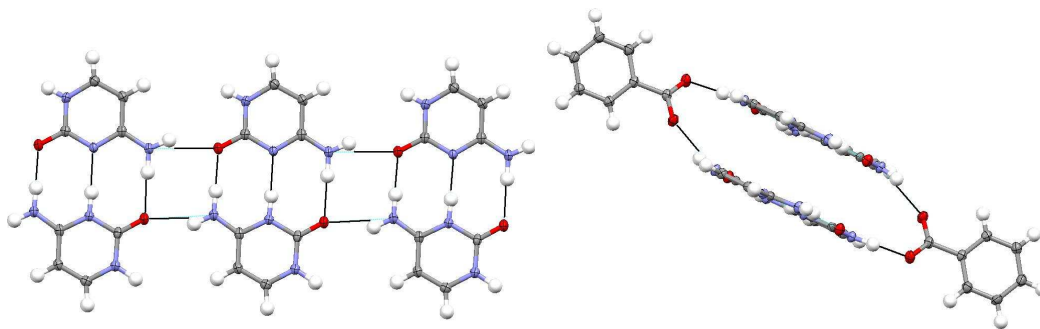


Figure 4.74 Left, hydrogen bonded chains and right, linking of the chains in the molecular complex of cytosine and benzoic acid

In the molecular complex formed between cytosine and benzoic acid, the pseudo-Watson Crick base pairing motif is also observed. These units are formed into flat chains which are bridged by benzoate molecules (Figure 4.74). However, the introduction of water molecules, the increased hydrogen bonding available via the hydroxy substituent on the benzoic acid, and the different crystallisation ratios give rise to significantly different extended structures. This increased hydrogen bonding stops the formation of extended cytosine chains in all three structures by termination by a benzoic acid, a water molecule or a combination of both.

The pseudo-Watson-Crick base pairs are formed in the cytosine and 5-fluorocytosine with 5-fluorosalicylic acid molecular complexes, but there is also a type B base pairing and the combination of these connects four cytosine or 5-fluorocytosine molecules together (two neutral and two charged) (Figure 4.75). The termination of the chain is through a combination of hydrogen bonding involving water molecules and 5-fluorosalicylate molecules (Figure 4.76). A moderate N-H...O hydrogen bond is formed from the protonated ring nitrogen to the carboxylate group of the 5-fluorosalicylate molecule and a weak C-H...O hydrogen bond from an aromatic ring hydrogen on the cytosine to the other oxygen of the carboxylate group (Table 4.17). The first water molecule forms two hydrogen bonds within the same plane; one O-H...O to the carbonyl oxygen of the cytosine of O...O distances of 2.838(2) Å and 2.857(2) Å for the cytosine and 5-fluorocytosine molecules, respectively; a second to the carboxyl group with O...O distances of 2.749(2) Å and 2.846(2) Å for the cytosine and 5-fluorocytosine molecules, respectively. A second water molecule connects layers of these units through hydrogen bonds to the amine groups of the cytosine molecules with N...O distances of 2.879(2) Å and 2.885(2) Å for the cytosine and 5-fluorocytosine molecules, respectively. The resultant structures are layered in nature, and the connection between layers is due to hydrogen bonds involving one of the water molecules and through base stacking.

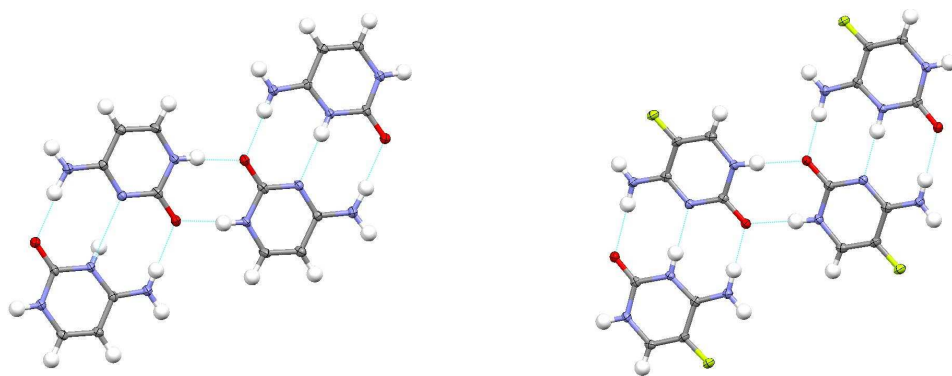


Figure 4.75. Pseudo-Watson-Crick and type B base-pairing in the cytosine and 5-fluorosalicylic acid molecular complex (left) and the 5-fluorocytosine and 5-fluorosalicylic acid molecular complex (right).

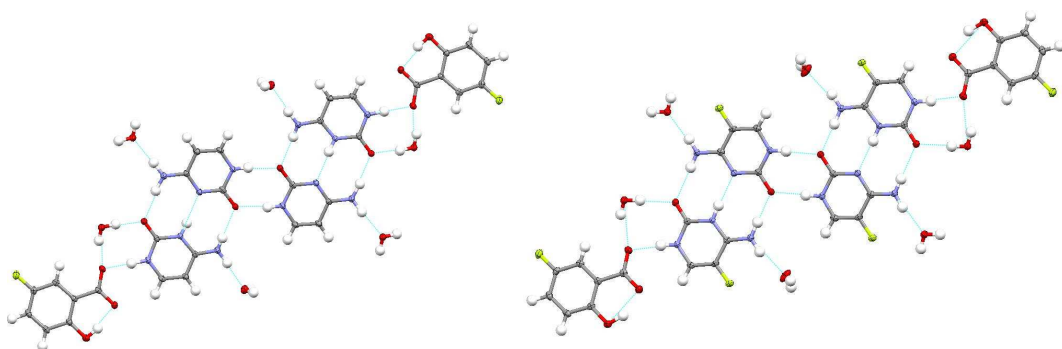


Figure 4.76 Termination of the chain in the cytosine and 5-fluorosalicylic acid molecular complex (left) and 5-fluorocytosine and 5-fluorosalicylic acid molecular complex (right)

Table 4.17. Hydrogen bonding in the isostructures of cytosine and 5-fluorosalicylic acid hydrate and 5-fluorocytosine and 5-fluorosalicylic acid hydrate

	Cytosine and 5-fluorosalicylic acid hydrate	5-fluorocytosine and 5-fluorosalicylic acid hydrate
N-H...O (protonated nitrogen to carboxylate group) (Å)	2.718(2)	2.684(2)
C-H...O (aromatic ring hydrogen to carboxylate group) (Å)	3.185(3)	3.097(2)
O-H...O (1 st water molecule and the carbonyl oxygen of the cytosine (Å)	2.838(2)	2.857(2)
O-H...O (1 st water molecule and the carboxyl group) (Å)	2.749(2)	2.846(2)
N-H...O (2 nd water molecule and the amine group) (Å)	2.879(2)	2.885(2)

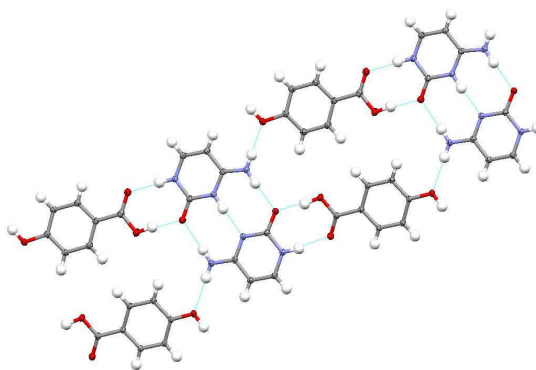


Figure 4.77. Extended hydrogen bonded chain of alternating cytosine and 4-hydroxybenzoic acid molecules in the cytosine and 4-hydroxybenzoic acid molecular complex

The cytosine 4-hydroxybenzoic acid molecular complex crystallises in an unusual 2:3 ratio of cytosine to 4-hydroxybenzoic acid. This leads to the presence of two neutral 4-hydroxybenzoic acid molecules and one 4-hydroxybenzoate molecule in the molecular complex in addition to the cytosine and cytosinium molecules (and water). The cytosine chain is terminated by 4-hydroxybenzoic acid molecules, however, these molecules become part of the chain spacing the base pairs, a consequence of the crystallisation ratio (Figure 4.77). A hydrogen-bonded ring is formed with the cytosine through the carboxyl group, the protonated ring nitrogen of the cytosine and the carbonyl oxygen of the cytosine. A single hydrogen bond is also formed between the amine group of the cytosine to hydroxy oxygen of the 4-hydroxybenzoic acid. The 4-hydroxybenzoate molecule, in combination with the water molecules does, however, still act as a bridging unit between layers of these cytosine 4-hydroxybenzoic acid chains.

In summary, the 4-hydroxybenzoic acid and cytosine structure is not as closely related to the other two structures due to the different basic building block. However, the 5-fluorosalicylic acid : 5-fluorocytosine hydrate and the cytosine : 5-fluorosalicylic acid structures are very closely related.

4.10.5. The Effect of the Fluorine

The presence of fluorine in the molecular complexes presented in this chapter does not interrupt any of the primary base-pairing motifs that are possible; the primary base-pairing motif is instead dependent on proton transfer and the crystallisation ratio. The most common interaction involving fluorine, which is found in all of the fluorinated molecular complexes, is a weak C-H...F hydrogen bond. Fluorine-fluorine interactions are also observed but these are much less common. These interactions have been found to play a subtle role in similar molecular complexes. This is clearly demonstrated by the presence of the additional fluorine atom in the 5-fluorosalicylic acid and 5-fluorocytosine hydrate

molecular complex compared to the 5-fluorosalicic acid and cytosine hydrate molecular complex; the additional fluorine atom is involved in a weak hydrogen bond and fluorine-fluorine interactions.

The fluorine atom also acts as a link between layers through fluorine-fluorine interactions. This type of interaction is seen in the 5-fluorocytosine and salicylic acid molecular complex and also the 5-fluorocytosine and 3-hydroxybenzoic acid complex.

4.10.6. Base Stacking Interactions

Base stacking interactions are found in all of the structures presented in this chapter. The majority of the structures contain a staggered stacking motif (Figure 4.78). The staggered stacking results in molecules lying over a hydrogen bonded pair - the layer spacing in these cases is shorter than that found where there is no staggering. The molecular complexes that do not show this staggering of the layers with respect to one another have a cytosine stacked on top of another cytosine. In the cases where there are mixed cytosine and cytosinium molecules, cytosinium molecules are stacked upon neutral cytosine molecules. In the hydrated molecular complexes, base stacking is still present but the strongest interactions between layers are hydrogen bonds involving the water molecules creating a stronger three dimensional hydrogen bonded structure with strong links between layers rather than the weaker base stacking interactions and fluorine interactions between the layers.

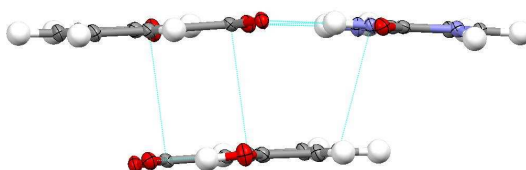


Figure 4.78. Staggered stacking interactions observed in several structures in this chapter.

Chapter 5 Molecular Complexes of Urea and Thiourea with Cytosine and Uracil and their Fluorinated Derivatives

The structures presented in this chapter are based around four base molecules (cytosine, 5-fluorocytosine, uracil and 5-fluorouracil) and the molecular complexes that they form with urea and thiourea.

5.1. 1:1 Molecular Complex of 5-Fluorouracil and Urea

5-Fluorouracil and urea were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless plates and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.5mm x 0.2mm x 0.1mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 5.8 at the end of this Chapter.

The molecular complex forms in a 1:1 ratio of 5-fluorouracil to urea. The pseudo Watson Crick base pair is not possible in the uracil molecular complexes as both heteroatoms of the ring are protonated. This structure contains no homo base pairing; there is a hetero base-pair motif between the uracil and the urea forming a hydrogen-bonded ring (Figure 5.1).

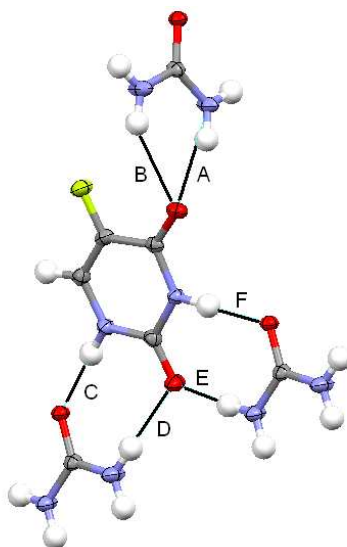


Figure 5.1. The hetero base pairing between 5-fluorouracil and urea in the 1:1 molecular complex.

Table 5.1. The hydrogen bonds involving the 5-fluorouracil molecule in the 1:1 molecular complex with urea. A-E represent the labelling in Figure 5.1.

Hydrogen Bond	D...A Distance (Å)	H...A Distance (Å)	DHA Angle (°)
A	3.011(3)	2.15(4)	153(4)
B	3.134(3)	2.39(5)	140(4)
C	2.779(3)	1.85(4)	175(3)
D	2.865(3)	1.98(4)	168(3)
E	2.852(3)	1.99(4)	167(3)
F	2.793(3)	1.80(4)	175(3)

The 5-fluorouracil molecule is involved in six hydrogen bonds (Table 5.1); there are two hetero base-pairing motifs involving the oxygen *para* to the fluorine and the heteroatoms. The oxygen acts as a hydrogen bond acceptor and forms two independent N-H...O hydrogen bonds to the amine groups of two symmetry related urea molecules; both are of moderate strength with N...O distances of 2.852(3) and 2.865(3) Å. The two heteroatoms in the 5-fluorouracil ring act as hydrogen bond donors to the oxygen atoms on the two urea molecules. These bonds are also of moderate strength and have N...O distances of 2.793(3) Å and 2.779(3) Å. These hetero base-pair motifs form into chains of alternating urea and 5-fluorouracil molecules (Figure 5.2). There is an additional weak C-H...O hydrogen bond between 5-fluorouracil molecules within the same chain. This has a C...O distance of 3.522 (2) Å.

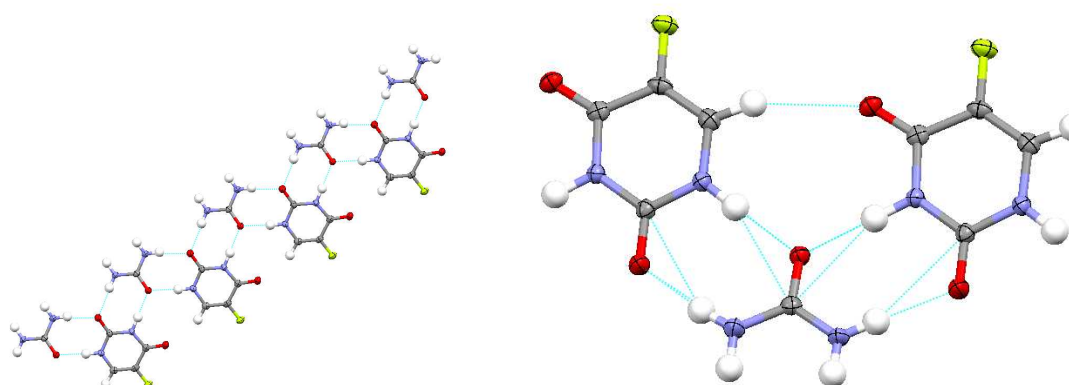


Figure 5.2. Hetero base- paired chains of urea and 5-fluorouracil.

Parallel chains are connected to one another through N-H...O bifurcated hydrogen bonding (Bond A and B, Figure 5.1) to chains which run almost perpendicular to them (Figure 5.3). These hydrogen bonds are of moderate strength and have N...O distances of 3.011(2) Å. This creates 10 molecule hydrogen bonded rings. The interlocking chains produce a net-like structure (Figure 5.4).

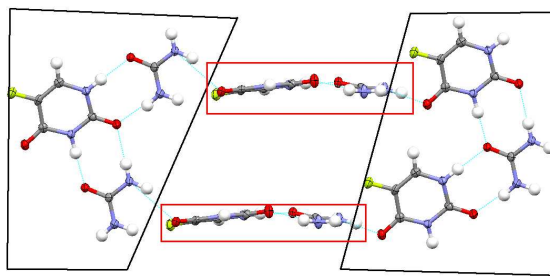


Figure 5.3. 10 membered ring created by hydrogen bonds in the molecular complex of 5-fluorouracil and urea.

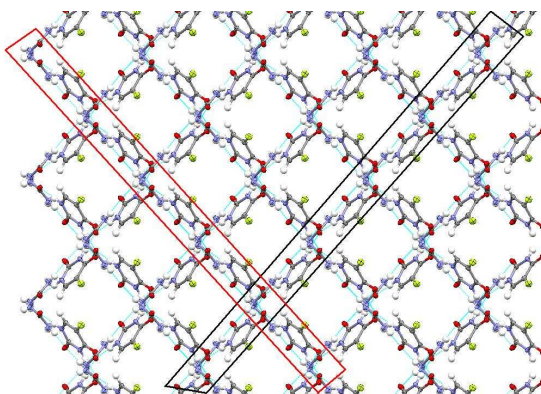


Figure 5.4 Net-like structure in the 5-fluorouracil and urea molecular complex. The black box represents the chain formed in Figure 5.3 and the red box represents one of the two chains in Figure 5.3.

A chain belonging to an inverted net-like structure is located in the centre of the ten molecule hydrogen bonded ring. This is held in place by a combination of fluorine-fluorine and base-stacking interactions (Figure 5.5, left).

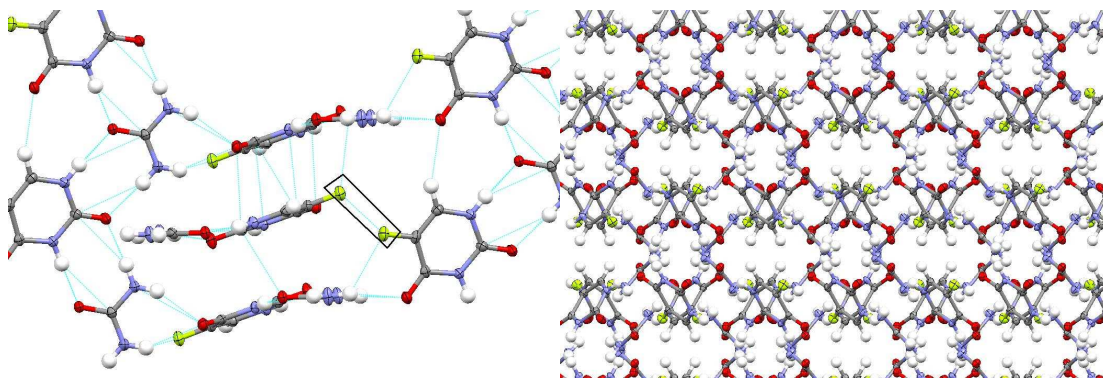


Figure 5.5. Left, fluorine=fluorine and base-stacking interactions holding the inverted chain in the centre of the ten molecule hydrogen bonded ring; right the crystal packing viewed along the c axis in the 5-fluorouracil and urea molecular complex.

The green and yellow chains in Figure 5.6 make up one of the net-like structures and the purple and white chains make up the other net-like structure; these are inversions of one another. The two structures are closely intertwined.

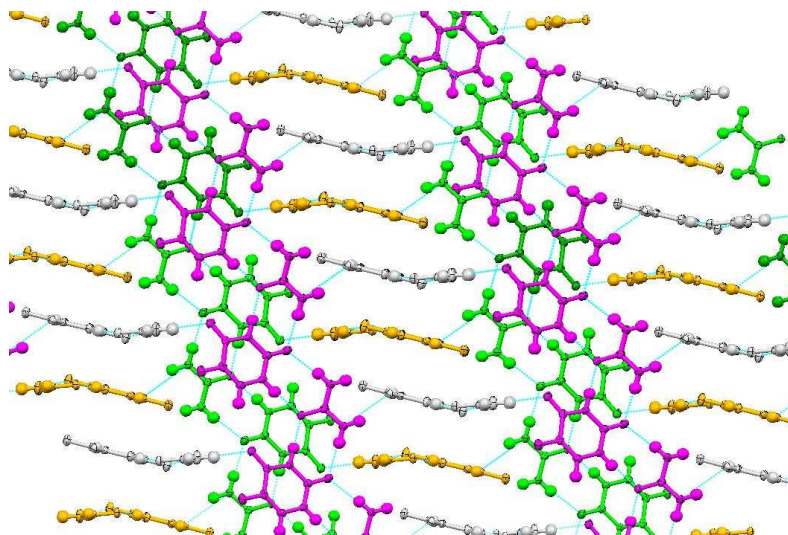


Figure 5.6. The intertwined nets in the molecular complex of 5-fluorouracil with urea. The first net is formed between the green and orange chains, and its inverted equivalent is between the white and purple.

There are no strong interactions between molecules of the same type in this molecular complex. There are fluorine-fluorine interactions between 5-fluorouracil molecules in approximately perpendicular chains of length 2.666(3) Å (Figure 5.7). This is significant when compared to the sum of the van der Waals radii for two fluorine atoms (2.94 Å). There is also an N-H...F weak hydrogen bond between the amine group of a urea molecule in a different almost perpendicular chain. This bond has an N...F distance of 3.218(3) Å.

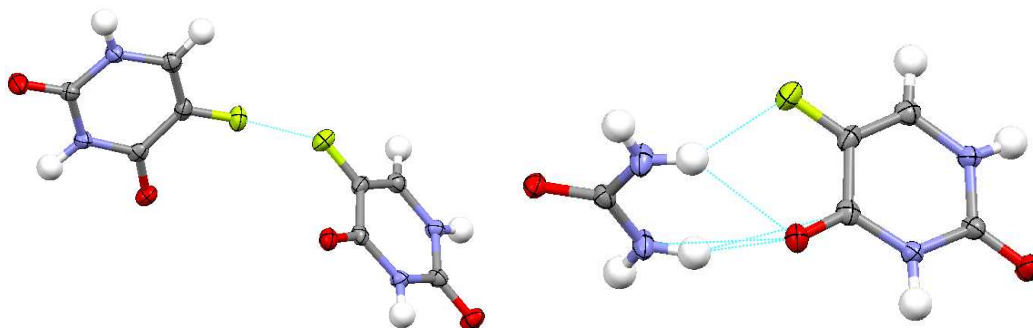


Figure 5.7. Fluorine interactions in the 5-fluorouracil and urea molecular complex.

One important factor that is present in the majority of structures involving any of the bases, fluorinated or natural, is base stacking. There are two distinctly different stacking interactions in this molecular complex (Figure 5.8). One involves a urea stacked above a 5-fluorouracil molecule and the other involves a 5-fluorouracil stacked above a separate 5-fluorouracil molecule. The 5-fluorouracil stacking interactions show a staggered configuration with the molecules being shifted by the width of half a 5-fluorouracil molecule. The stacking interaction here was calculated by calculating a best plane for each 5-fluorouracil molecule involved, then measuring the perpendicular difference between these two planes; in this case this is approximately 3.095 Å.

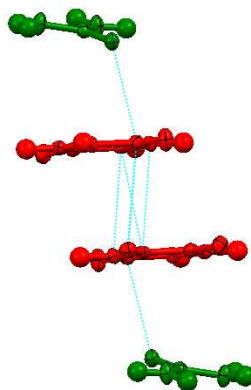


Figure 5.8. Stacking interactions in the 5-fluorouracil and urea molecular complex.

The other stacking interaction involves a urea molecule stacked on top of a 5-fluorouracil molecule, however, as these two planes are not parallel with respect to each other the distance between the centroid of the π bond on the 5-fluorouracil and the oxygen atom of the urea molecule was used to estimate the stacking distance; this was approximately 3.002 Å.

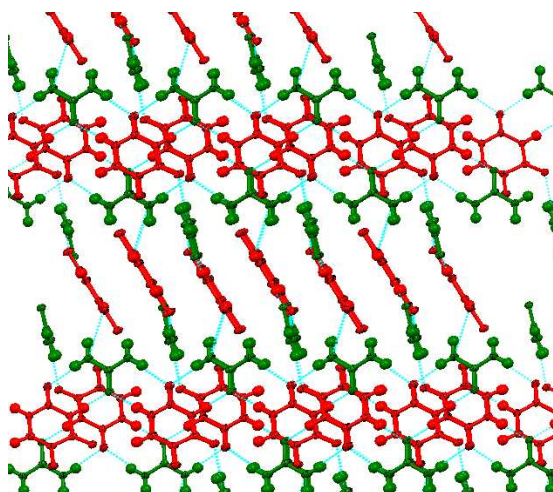


Figure 5.9. Global packing in the 5-fluorouracil and urea molecular complex. The 5-fluorouracil molecules are shown in red and the urea molecules in green.

The global packing of the structure shows that there are defined regions or “rows” of like molecules. Figure 5.9 shows that the “row” of 5-fluorouracil molecules has these all positioned in the same orientation (red) whilst the “row” of urea molecules has every second molecule positioned in a perpendicular manner to the previous one. This shows the role the urea molecule has in linking the structure three dimensionally through hydrogen bonding.

5.2. 1:1 Molecular Complex of Uracil and Urea

Uracil and urea were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless plates and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.5mm x 0.2mm x 0.1mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 5.8.

The molecular complex forms in a 1:1 ratio of uracil to urea and contains no homo-base-pairing but there is a hetero base paired motif between the uracil and the urea (Figure 5.10). This molecular complex is isostructural with the urea and 5-fluorouracil molecular complex.

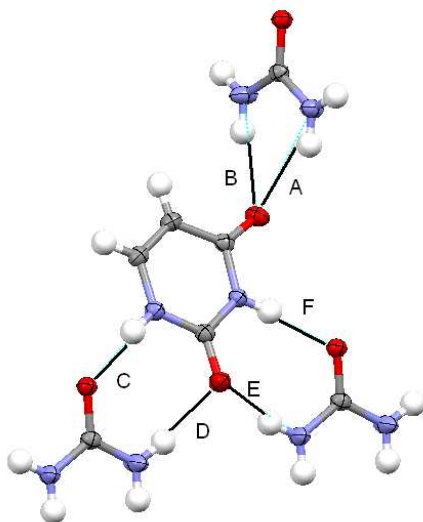


Figure 5.10. The hydrogen bonding environment of the uracil molecule in its molecular complex with urea.

Table 5.2. The hydrogen bonds involving the 5-fluorouracil molecule in the 1:1 molecular complex with urea. A-E represent the labelling in Figure 5.10.

Hydrogen Bond	D...A Distance (Å)	H...A Distance (Å)	DHA Angle (°)
A	3.001(4)	2.25(5)	146(4)
B	2.927(4)	2.14(4)	150(3)
C	2.781(4)	1.93(5)	172(4)
D	2.873(4)	2.02(5)	170(4)
E	2.883(4)	2.06(4)	159(3)
F	2.794(4)	1.94(4)	174(4)

This structure, as seen in the equivalent 5-fluorouracil complex, has the uracil molecule involved in six hydrogen bonds (Table 5.2); there are two hetero base-pairing motifs

involving the oxygen positioned between the two heteroatoms of the cytosine ring. The oxygen atom acts as a hydrogen bond acceptor and forms two independent N-H...O hydrogen bonds to the amine groups of two symmetry related urea molecules; both are of moderate strength with N...O distances of 2.873(4) Å and 2.883(4) Å. To complete the hetero base-pairing motifs, the two heteroatoms in the uracil ring act as hydrogen bond donors to the oxygen atoms on the two urea molecules. These bonds are also of moderate strength and have N...O distances of 2.794(4) Å and 2.781(4) Å. These form into hetero base-paired chains of alternating urea and uracil molecules (Figure 5.11).

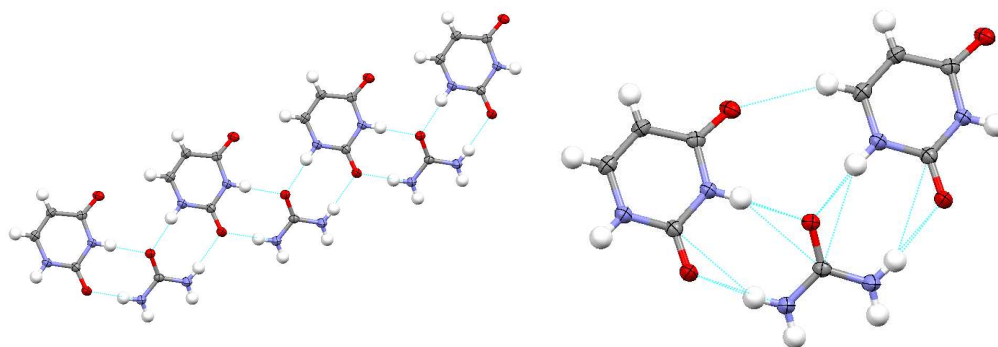


Figure 5.11. Hetero base- paired chains of urea and uracil.

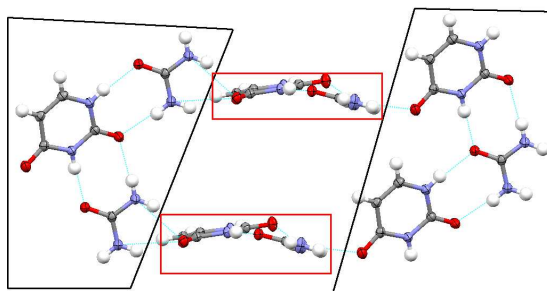


Figure 5.12. 10 membered ring created by hydrogen bonds in the molecular complex of uracil and urea.

Figure 5.12 shows how the parallel chains are connected to one another through N-H...O bifurcated hydrogen bonding (Bond A and B, Figure 5.10) to chains which run almost perpendicular to them (Figure 5.12). These hydrogen bonds are of moderate strength and have N...O distances of 3.001(4) Å and 2.927(4) Å. This creates 10 molecule hydrogen bonded rings. The interlocking chains produce a net-like structure (Figure 5.13).

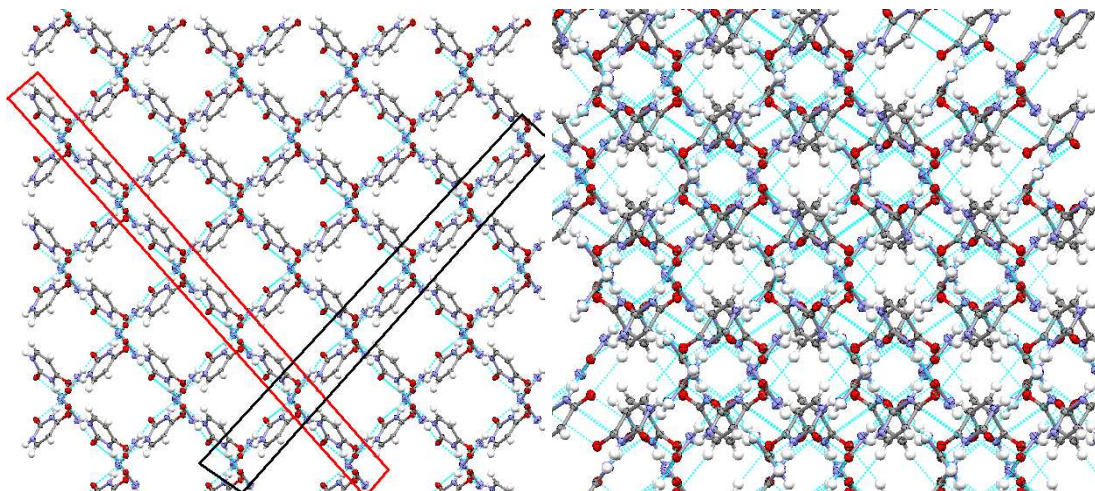


Figure 5.13. Left, the net like structure and right, the crystal packing of intertwined nets viewed along the c axis in the uracil and urea molecular complex.

A chain belonging to an inverted net-like structure is located in the centre of the ten molecule hydrogen bonded ring (Figure 5.14). However this is, for obvious reasons, not held in place by fluorine interactions but solely by base-stacking interactions.

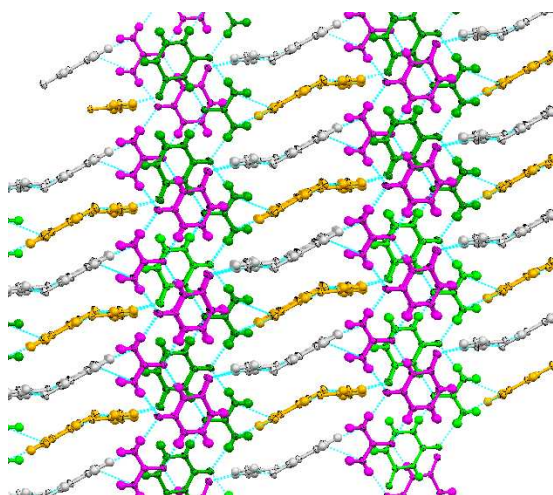


Figure 5.14. The intertwined nets in the molecular complex of uracil with urea. The first net is formed between the green and orange chains, and its inverted equivalent is between the white and purple.

With no fluorine atoms, the main link between layers, aside from the previously described hydrogen bonds, is the stacking interactions. The stacking interactions are very similar to those in the previous section in both the positioning and the arrangement. Once again there are two distinct stacking interactions in this structure, one between two uracil molecules and one between a uracil and a urea molecule (Figure 5.15). The distances between the planes were calculated in the same manner as the 5-fluorouracil and urea structure. The distance between the two uracil molecules is approximately 3.160 Å and the distance between the uracil and urea molecules is approximately 3.091 Å.

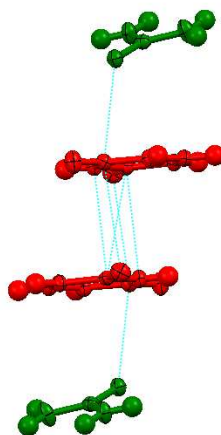


Figure 5.15. Stacking interactions in the uracil and urea molecular complex.

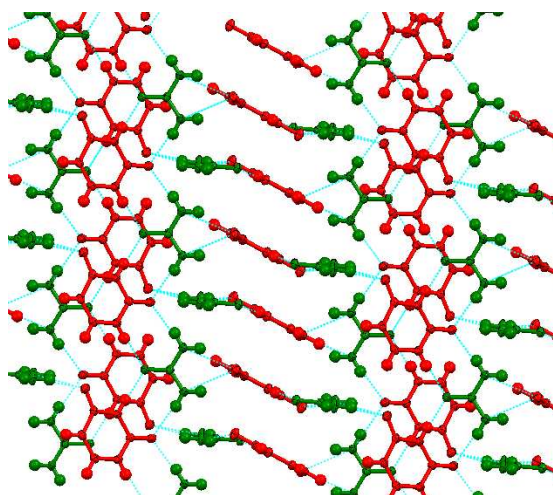


Figure 5.16. Global packing in the uracil and urea molecular complex. The uracil molecules are shown in red and the urea molecules in green.

The global packing, seen in this structure, shows that there are defined regions or “rows” of like molecules. The “row” of uracil molecules are all positioned in the same orientation (red in Figure 5.16), whilst the “row” of urea molecules has every second molecule positioned in a perpendicular manner to the last one (green in Figure 5.16). Urea molecules again link the structure three dimensionally through hydrogen bonding.

5.3 1:1 Molecular Complex of 5-Fluorouracil and Thiourea

5-Fluorouracil and thiourea were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless plates and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.5mm x 0.2mm x 0.1mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 5.8.

The molecular complex forms in a 1:1 ratio of 5-fluorouracil to thiourea. There is homo base pairing in this structure (Figure 5.17) and also a hetero-pseudo-base-pairing motif between the 5-fluorouracil and thiourea molecules.

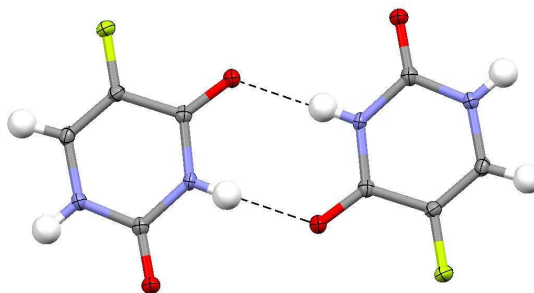


Figure 5.17. Homo base-pairing in the 5-fluorouracil and thiourea molecular complex.

The homo base pair in this structure is a pseudo Hoogsteen of type B, which involves two moderate strength hydrogen bonds with both being of equivalent length of 2.808(2) Å; there is an inversion centre in the centre of the hydrogen bonded ring ($R_2^2(8)$). Unlike other similar structures, a homo base paired chain does not form in this case. The chain here involves both 5-fluorouracil and thiourea molecules where a 5-fluorouracil dimer is spaced by two thiourea molecules (Figure 5.18).

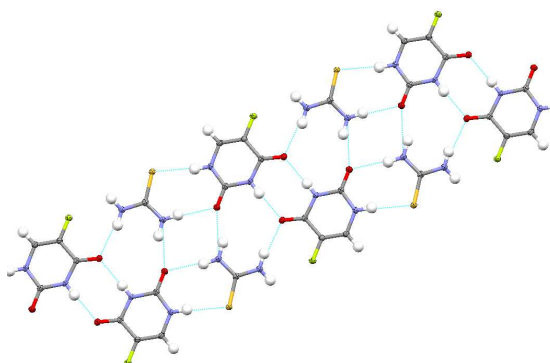


Figure 5.18. Alternating pairs of 5-fluorouracil and thiourea molecules in the hydrogen bonded chain. The hetero pseudo-base-pairs that form between a thiourea and a 5-fluorouracil molecule consist of N-H...S and a N-H...O hydrogen bonds of N...S and N...O distances of 3.184(3) Å (D in Figure 5.19) and 3.024(3) Å (E in Figure 5.19), respectively. The linking of the two types of base pairing is shown in Figure 5.20. In addition, there are two further hydrogen bonded rings formed involving both thiourea and 5-fluorouracil. The first involves two 5-fluorouracil molecules and one thiourea molecule with three N-H...O hydrogen bonds (A, B and C in Figure 5.19). This ring has graph set notation of $R_4^2(8)$. The second forms a diamond shape and involves two 5-fluorouracil molecules and two thiourea molecules with four N-H...O hydrogen bonds (alternating C and E Figure 5.19)

with an inversion centre in the centre of the ring. This ring also has graph set notation of $R_4^2(8)$.

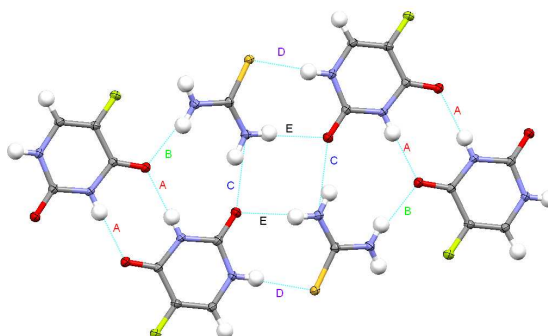


Figure 5.19. The hydrogen bonding pattern in the chains of alternating 5-fluorouracil and thiourea molecules.

Table 5.3. The hydrogen bonds involving the 5-fluorouracil and thiourea molecules in the 1:1 molecular complex of 5fluorouracil and thiourea. A-E represent the labelling in Figure 5.19.

Hydrogen Bond	D...A Distance (Å)	H...A Distance (Å)	DHA Angle (°)
A	2.808(2)	1.99(2)	176(2)
B	2.922(2)	2.06(2)	167(2)
C	2.911(2)	2.05(2)	154(2)
D	3.184(1)	2.30(2)	174(2)
E	3.023(2)	2.17(2)	168(2)

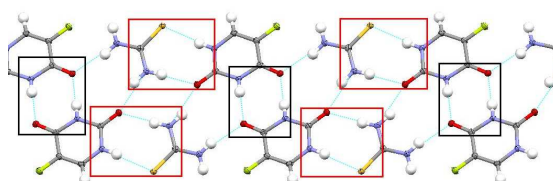


Figure 5.20. The two types of base-pairing in the molecular complex of 5-fluorouracil with thiourea. Black indicates the homo base pair and red represents the thiourea 5-fluorouracil hetero base pair motif.

In this structure, the fluorine atom on the 5-fluorouracil molecule is involved in a weak hydrogen bond (Figure 5.21). This is the only interaction in which the fluorine is involved. This interaction involves one of the hydrogen atoms of one of the amine groups of the thiourea and the fluorine on the 5-fluorouracil. This bond is of moderate strength and has an N...F distance of 2.931(3) Å. These interactions supplement the other hydrogen-bonded interactions of each chain.

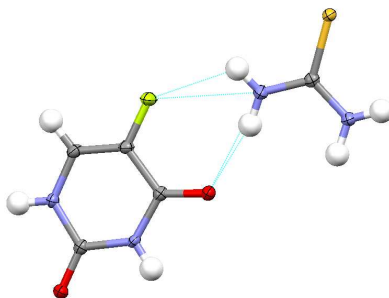


Figure 5.21. Weak hydrogen bond involving the fluorine atom in the 5-fluorouracil and thiourea molecular complex.

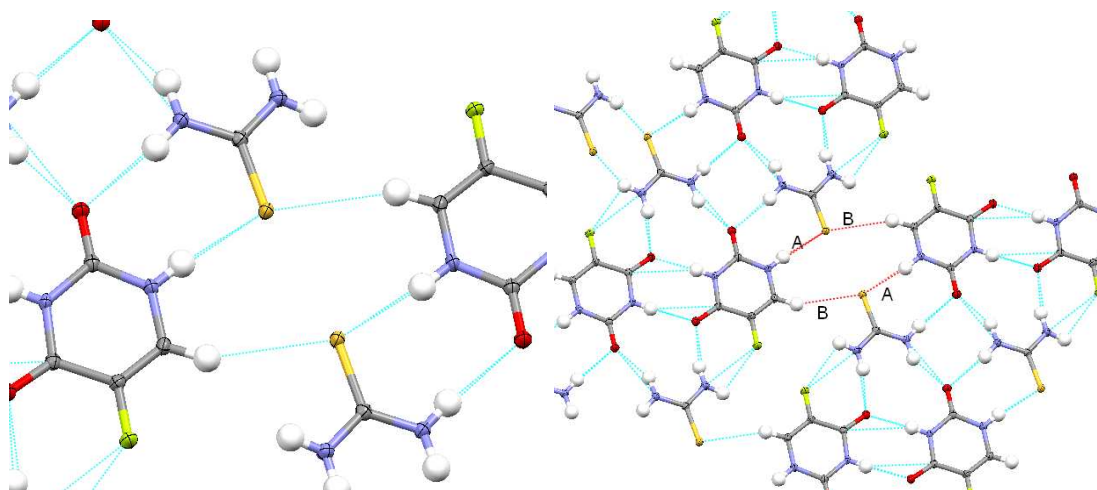


Figure 5.22. Weak hydrogen bond linking chains of 5-fluorouracil and thiourea.

Weak hydrogen bonds provide a link between chains creating layers. A ring made up of four hydrogen bonds and two unique distances present is formed (Figure 5.22). The first of these hydrogen bonds is a weak N-H...S hydrogen bond, from the hydrogen on a heteroatom next to the oxygen in the uracil ring and the sulphur of a nearby thiourea. This N...S bond distance is 3.184(2) Å and is of moderate strength. This bond appears twice in the ring formed (Figure 5.22 right, A). The second hydrogen bond present is a C-H...S hydrogen bond and involves the hydrogen on a carbon in between the fluorine atom and a heteroatom in the ring and the same sulphur atom involved in the first hydrogen bond. This bond is of weak strength and has a C...S distance of 3.720(2) Å. These hydrogen bonds combine to link the previously described chains together.

There is a weak N-H...S hydrogen bond between adjacent layers where the amino group of the thiourea molecule is slightly twisted out of the plane (Figure 5.23). This twisting may be what causes the chains, described earlier, to deviate from planarity as it is clear to see that the thiourea molecules are twisted out the plane towards each other. This interaction is made up of two equivalent hydrogen bonds. This forms what resembles a base-paired motif with two hydrogen bonds forming between equivalent areas of the two molecules involved

in the bonding motif. The bonds are from the amine group of one thiourea to the sulphur of a nearby thiourea. The bonds are of weak strength with an N...S distance of 3.439(2) Å.

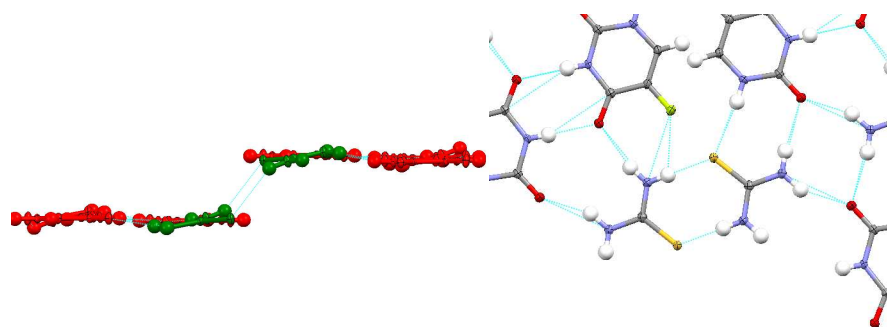


Figure 5.23. Sulphur interactions between the layers in the molecular complex of 5-fluorouracil and thiourea.

There are also π - π interactions between parallel chains (Figure 5.24). Structures that involve bases usually have base-stacking interactions, however, here the stacking interactions are staggered to an extent that the 5-fluorouracil is stacked on top of the thiourea (Figure 5.24), meaning that the π - π interactions are between the 5-fluorouracil and thiourea. The approximate stacking distance is 3.281 Å.

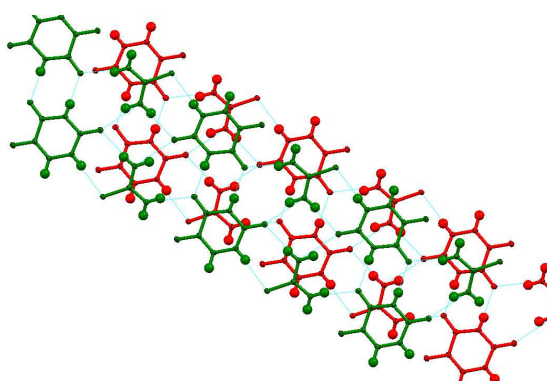


Figure 5.24 Overlaid chains in the molecular complex of 5-fluorouracil and thiourea.

5.4 3:1:1 Molecular Complex of 5-Fluorocytosine Thiourea Hydrate

5-Fluorocytosine and thiourea were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless plates and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.3mm x 0.3mm x 0.1mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 5.8.

The molecular complex forms in a 3:1:1 ratio of 5-fluorocytosine to thiourea to water. This structure contains no hydrogen transfer and contains three independent 5-fluorocytosine molecules. For structures that contain no hydrogen transfer the most common bonding motifs adopted are the type A and type B and in many structures these will combine to

form elongated chains. This structure does contain these elongated chains; in fact, there are two independent chains of this type. These chains both contain the same combination of type A and type B bonding motifs, but they are different in that one is made up of two independent 5-fluorocytosine molecules (chain 2) whilst the other chain is made up of one independent 5-fluorocytosine molecule (chain 1).

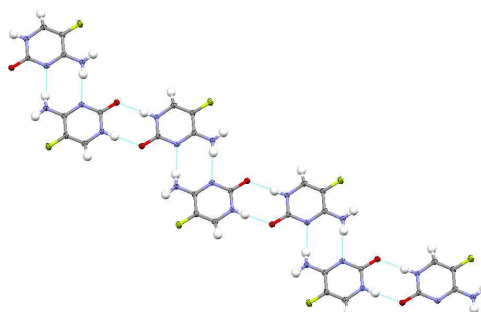


Figure 5.25. Extended 5-fluorocytosine chain (chain 1) formed in the molecular complex of 5-fluorocytosine and thiourea hydrate.

In chain 1, the type A bonding motif is made up of two equivalent N-H...N bonds (Figure 5.25). These bonds involve a hydrogen atom of the amine group of the 5-fluorocytosine and the unprotonated heteroatom of the ring of a nearby 5-fluorocytosine. This bond is reciprocated by inversion symmetry to produce two equivalent hydrogen bonds. These bonds are of moderate strength and have an N...N distance of 3.009(2) Å.

The other bonding motif in this chain is the type B bonding motif and this is made up of two equivalent N-H...O hydrogen bonds. These involve the hydrogen on the protonated heteroatom of the 5-fluorocytosine ring and the oxygen atom of a nearby 5-fluorocytosine. Like the bonding present in the type A bonding motif, this bond is reciprocated by inversion to produce two equivalent hydrogen. The hydrogen bonds are therefore both of moderate strength and have an N...O distance of 2.784(2) Å. All of the 5-fluorocytosine molecules in this chain are equivalent.

The same can not be said of chain 2 which is constructed of the two remaining independent 5-fluorocytosine molecules present in the structure. The same type of chain is formed with the alternating type A and B bonding motifs being deployed to form an extended chain. However, only every second molecule is equivalent as shown in Figure 5.26.

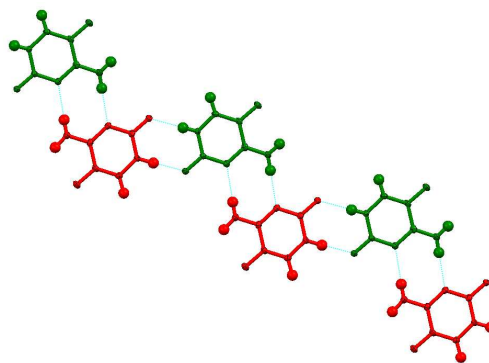


Figure 5.26. The two independent 5-fluorocytosine molecules (one green and one red) combine in the molecular complex of 5-fluorocytosine and thiourea hydrate to form an extended chain (chain 2).

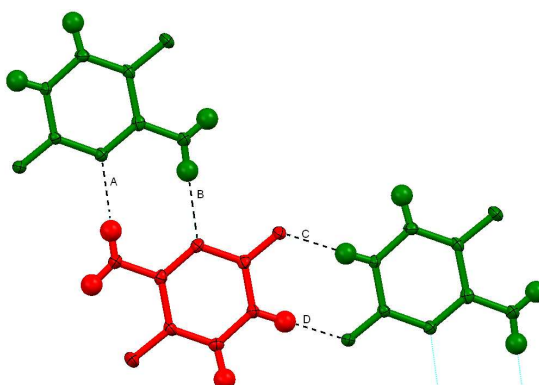


Figure 5.27. All hydrogen bonds (labelled) in the 5-fluorocytosine chain 2 in the molecular complex of 5-fluorocytosine and thiourea hydrate.

The type A bonding motif is made up of two N-H...N hydrogen bonds in the same manner as the other chain, however, there are two distances here. The hydrogen bonds have N...N distances of (A) 2.958(3) Å and (B) 2.931(2) Å with both bonds being of moderate strength. The type B bonding motif is made up of two independent N-H...O hydrogen bonds. These bonds have an N...O distance of (C) 2.750(2) Å and (D) 2.741(2) Å.

The thiourea molecule becomes involved with chain 2 via a bifurcated hydrogen bond. This bifurcated hydrogen bonding only occurs on one side of the chain and with only one of the independent molecules in the chain (Figure 5.28, left). The two hydrogen bonds in the bifurcated hydrogen bond involve two amine groups of the thiourea molecule and an oxygen atom of the 5-fluorocytosine molecule. The two hydrogen bonds are of moderate strength and have N...O distances of 2.956(3) Å and 2.896(2) Å. The thiourea molecule is orientated perpendicular to the chain. On the other side of the chain there is a thiourea tied to the chain via a single hydrogen bond to the oxygen atom of a 5-fluorocytosine molecule in the chain that is not involved in the bifurcated bond. This hydrogen bond involves just one of the amine groups of the thiourea and the oxygen atom of the 5-fluorocytosine molecule. The thiourea lies below the original plane and is orientated perpendicular to the original chain. This thiourea that is below the original plane is then involved in the

bifurcated bonding motif and therefore bridges two 5-fluorocytosine chains (chain 2) (Figure 5.29).

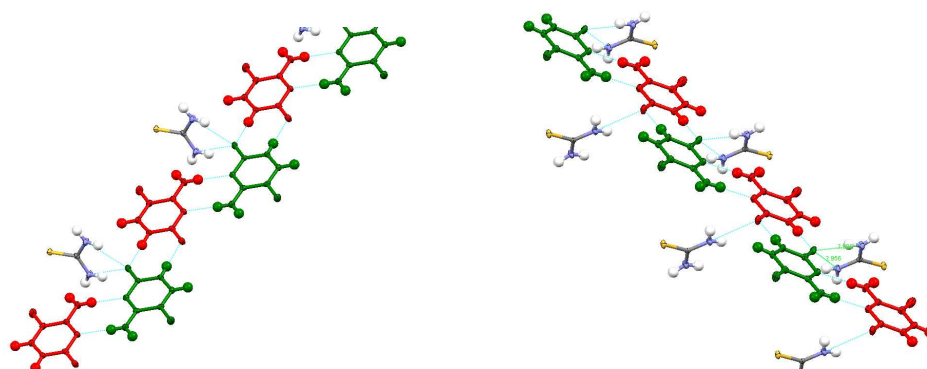


Figure 5.28. Thiourea linked to either side of the 5-fluorocytosine chain (chain 2) in the molecular complex of 5-fluorocytosine and thiourea hydrate.

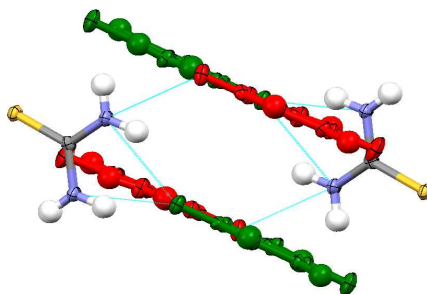


Figure 5.29. Thiourea linking separate layers in the molecular complex of 5-fluorocytosine and thiourea hydrate.

As a consequence there are base stacking interactions between the two chains bridged by the two thiourea molecules. These stacking interactions are not staggered and have a 5-fluorocytosine stacked in the same orientation on top of another 5-fluorocytosine molecule (Figure 5.30). However, these two 5-fluorocytosine molecules are not equivalent and two independent molecules are stacked on top of one another in this structure. The distances between the chains was estimated by calculating a centroid in the centre of the ring for each 5-fluorocytosine molecule in each chain, and then measuring the distance between them. The distance between the planes in this area of the structure is 3.439 Å.

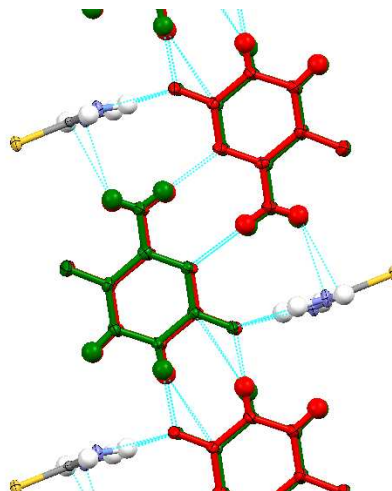


Figure 5.30. Chains overlayed in the molecular complex of 5-fluorocytosine and thiourea hydrate.

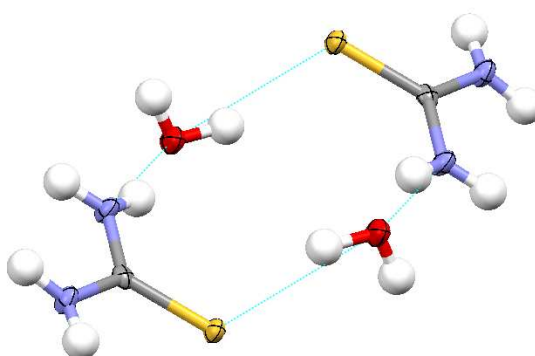


Figure 5.31. Hydrogen bonded ring formed between two thiourea molecules and two water molecules in the molecular complex of 5-fluorocytosine and thiourea hydrate.

The water molecules link thiourea molecules together via two unique hydrogen bonds, which form a ring comprised of two water and two thiourea molecules (Figure 5.31). The first hydrogen bond comes from the amine group of the thiourea and involves the oxygen atom of the water molecule to form a N-H...O hydrogen bond which is of moderate strength. The N...O distance is 2.859(2) Å. The second hydrogen bond involves a hydrogen atom of the water molecule and the sulphur of the thiourea. This hydrogen bond is of weak strength and has an O...S distance of 3.254(2) Å.

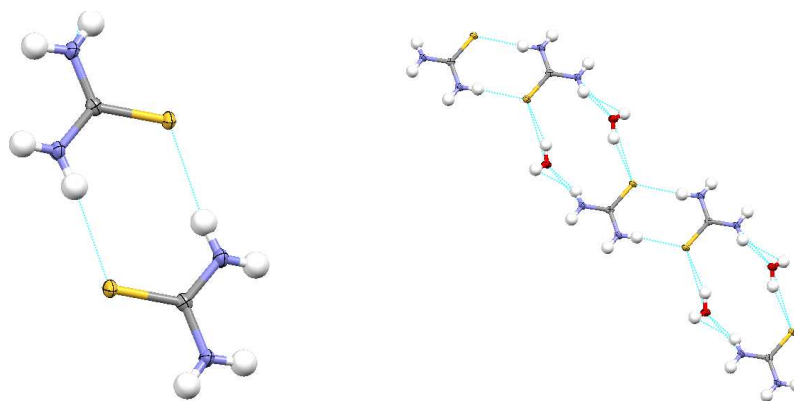


Figure 5.32. Left, hydrogen bonded ring formed between two thiourea molecules. Right, the hydrogen bonded ring between two thiourea molecules linking four membered rings together in the molecular complex of 5-fluorocytosine and thiourea hydrate.

This is not the only way of linking two thiourea molecules together in this structure. The other way is through weak N-H...S hydrogen bonds. This forms a bonding motif that resembles a base-paired motif. These bonds come from the amine group of one thiourea and goes to the sulphur atom of a nearby thiourea. This bond is reciprocated by inversion symmetry to produce the bonding motif. The weak hydrogen bonds here are both of length 3.365(2) Å. This bonding motif combines with the ring formed with the water molecules to form a chain (Figure 5.32).

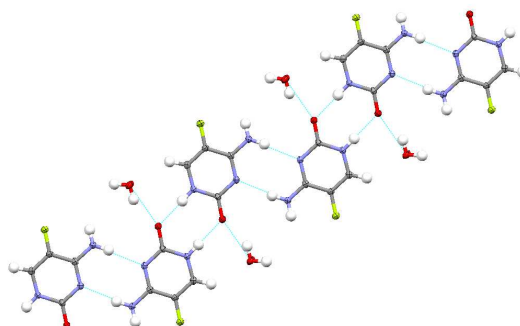


Figure 5.33. Water molecules tied to either side of the 5-fluorocytosine chain (chain 1) in the molecular complex of 5-fluorocytosine and thiourea hydrate.

The 5-fluorocytosine chain 1 which is comprised of only one independent 5-fluorocytosine molecule, is linked through a single hydrogen bond from the water molecule (Figure 5.33). The bond here is an O-H...O hydrogen bond of moderate strength and has an O...O distance of 2.723(2) Å.

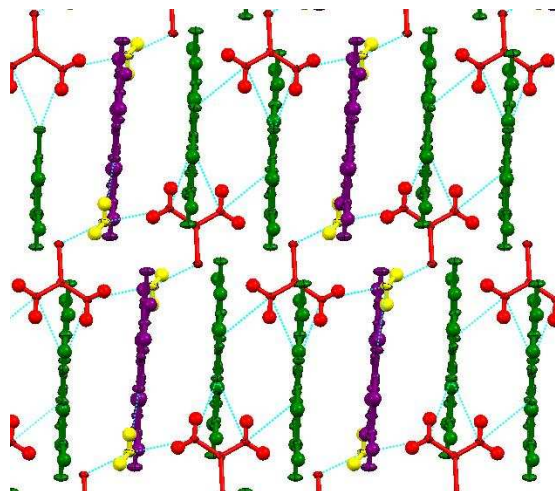


Figure 5.34. Water molecules (yellow) linking the separate layers together and the two independent chains in the structure (chain 1 purple and chain 2 green) in the molecular complex of 5-fluorocytosine and thiourea hydrate.

Figure 5.34 shows that the chains comprised of the two independent molecules (chain 2, green) are stacked on top of each other and the chain formed by just one independent molecule is stacked on top of this again (chain 1, purple). There are therefore two stacking interactions between chain 1 and chain 2 and these have approximate spacings of 3.307 Å and 3.254 Å.

The three independent 5-fluorocytosine molecules provide a wide opportunity for fluorine interactions. The fluorine atoms in chain 2 are both involved in very similar fluorine interactions. These interactions help to link separate layers together (Figure 5.35, right). There are two unique weak bonding motifs present. The first motif involves interactions A and B as seen in Figure 5.35, left. These bonds involve the fluorine of one 5-fluorocytosine molecule and the C-H group present in the ring or an adjacent 5-fluorocytosine. This bond is reciprocated to create a two hydrogen bond motif with two F...H-C hydrogen bonds of length 3.281(2) Å and 3.259(1) Å which are of weak strength. The second bonding motif is made up of the bonds involving the same areas of the 5-fluorocytosine molecules, but this motif is sandwiched by two occurrences of the first motif described. This motif is also made up of two F...H-C hydrogen bonds which are of weak strength and are both of length 3.182(1) Å.

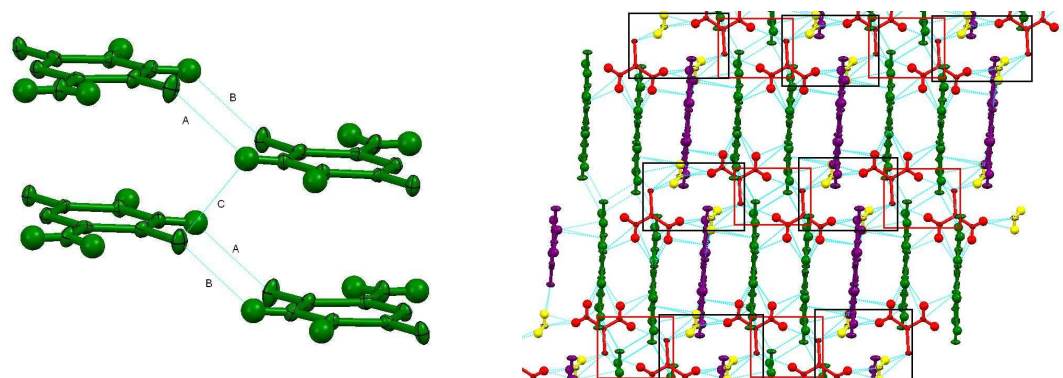


Figure 5.35. Left, fluorine interactions between layers and right, the fluorine interactions (red box) and the four membered ring of thiourea and water linking separate layers in the molecular complex of 5-fluorocytosine and thiourea hydrate.

5.5 2:1 Molecular Complex of 5-Fluorocytosine and Urea

5-Fluorocytosine and urea were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a plate like shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.2mm x 0.2mm x 0.1mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 5.8.

The complex forms in a 2:1 ratio of 5-fluorocytosine to urea. Base-paired chains of 5-fluorocytosine are formed between alternating independent 5-fluorocytosine molecules of pseudo-Hoogsteen type A and type B hydrogen bonded rings. The amine group hydrogen bonds with the unprotonated ring nitrogen and this bond is repeated in reverse to complete the base pair (C and D in Figure 5.36). The other base pair involves the hydrogen of the protonated nitrogen of the 5-fluorocytosine ring and the oxygen of another 5-fluorocytosine. This bond is repeated in reverse to complete the base pair (A and B in Figure 5.36).

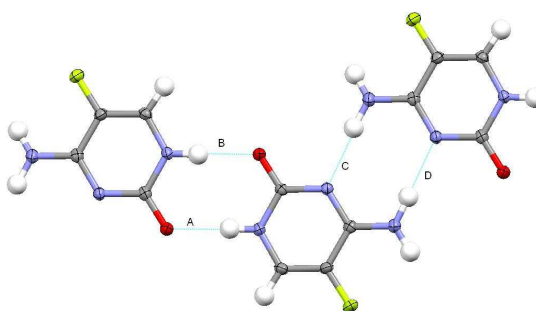


Figure 5.36. The base paired chains in the molecular complex of 5-fluorocytosine and urea.

Table 5.4. The hydrogen bonds involving the 5-fluorocytosine molecule in the 2:1 molecular complex with urea. A-D represent the labelling in Figure 5.36.

Hydrogen Bond	D...A Distance
A	2.757(3) Å
B	2.746(3) Å
C	2.960(2) Å
D	2.938(3) Å

The urea molecules are involved in a chain of their own forming a base-pairing type motif. The hydrogen-bonded ring is made up of two N-H...O hydrogen bonds, the first of which is from the oxygen atom of one urea molecule and the amine group of the next urea in the chain and this bond is reciprocated to complete the hydrogen-bonded ring. The two hydrogen bonds in each base pair are equivalent, there are two independent base pairing motifs (A and B in Figure 5.37). The hydrogen bonds making up bonding motif A are both of moderate strength and have N...O distances of 2.990(2) Å, whilst the bonds comprising bonding motif B both have an N...O distance of 2.933(2) Å. Bonding motifs A and B alternate as the chain is perpendicular to the 5-fluorocytosine chain.

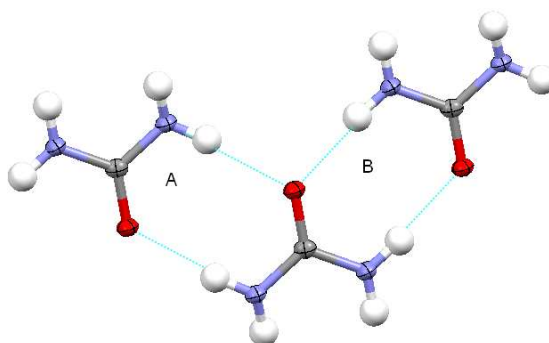


Figure 5.37. Urea hydrogen bonded chains in the molecular complex of 5-fluorocytosine and urea.

Separate chains of 5-fluorocytosine are linked via weak hydrogen bonds involving the fluorine atom on the 5-fluorocytosine. These resemble a base-paired motif with two weak C-H...F hydrogen bonds making up each connection between the chains (Figure 5.38) with C...F distances of 3.221(2) Å and 3.215(2) Å. This forms a 5-fluorocytosine framework.

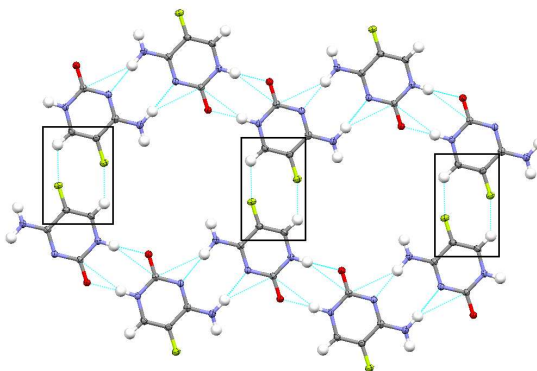


Figure 5.38. Fluorine interactions linking separate chains in the molecular complex of 5-fluorocytosine and urea.

The urea chain fits into the centre of the 5-fluorocytosine framework. On one side of the urea chain the urea is involved in a hydrogen bond with the amine group of a 5-fluorocytosine to create a moderate strength N-H...O hydrogen bond with an N...O distance of 2.952(2) Å. The other side of the urea chain is involved in two hydrogen bonds with the next 5-fluorocytosine chain. These two hydrogen bonds involve the same oxygen atom on a 5-fluorocytosine molecule and the two amine groups of the urea molecule. This creates a bifurcated hydrogen bond with two N-H...O hydrogen bonds of moderate strength and with N...O distances of 2.987(2) Å and 2.965(2) Å.

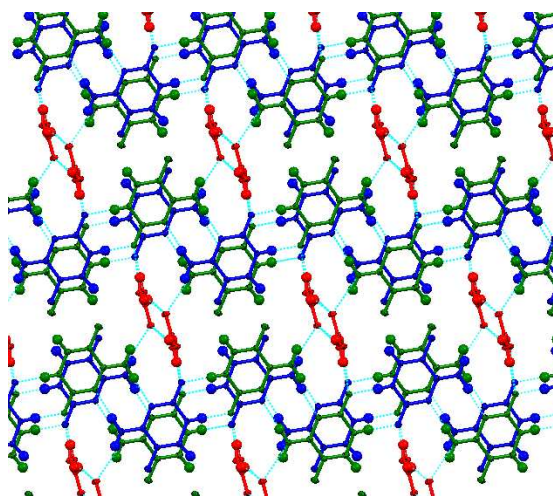


Figure 5.39. Colour coded structure with the cytosine molecules labelled blue and green and the urea molecules labelled red showing the positioning of the urea chains in relation to the 5-fluorocytosine chains in the molecular complex of 5-fluorocytosine and urea

This structure is very ordered with the 5-fluorocytosine chains forming with urea chains present at ordered intervals in the structure running perpendicular to the 5-fluorocytosine chains. The structure once again has base-stacking interactions present between the separate chains of 5-fluorocytosine. As can be seen from Figure 5.39 the chains line up closely but are slightly staggered. The base-stacking interactions were calculated by calculating centroids that lined up well between the layers as the layers are not perfectly planar. The distance between layers with the blue molecule on top of a green molecule was

found to be approximately 3.221 Å and with a green molecule on top of a blue molecule to be approximately 3.307 Å.

5.6 Comparison of Structures

Many co-crystallisations were set up with the aim to get a complete series of structures through the combination of uracil, 5-fluorouracil, cytosine and 5-fluorocytosine with both urea and thiourea. However, this proved to be problematic due to solubility issues. Water and methanol are the only solvents in which cytosine and 5-fluorocytosine are freely soluble and this restricts the crystallisation setup. A wide variety of techniques were employed (slow evaporation, solvent diffusion, vapour diffusion and all were carried out at a variety of temperatures), however, only complexes of 5-fluorouracil and urea, uracil and urea, 5-fluorouracil and thiourea, 5-fluorocytosine and urea, and 5-fluorocytosine and thiourea were obtained. In general, the complexes containing uracil and 5-fluorouracil all form in a 1:1 ratio. However, the complexes formed with 5-fluorocytosine show ratios of 2:1 and a 3:1:1 hydrate (Table 5.5).

Table 5.5. A summary of the ratios obtained in the molecular complexes involving urea and thiourea.

Structure	Crystallisation Ratio
5-Fluorouracil and Urea	1:1
Uracil and Urea	1:1
5-Fluorouracil and Thiourea	1:1
5-Fluorocytosine and Urea	2:1
5-Fluorocytosine and Thiourea Hydrate	3:1:1

Of these three uracil based molecular complexes, some common bonding patterns can be identified; the effect of exchanging thiourea and urea is more significant than fluorinating the uracil molecule. The molecular complexes of uracil and urea and 5-fluorouracil and urea are isostructural and the influence of the fluorine atom is far more subtle than that observed in the molecular complexes reported in other chapters.

5.6.1 Isostructural molecular complexes of 5-fluorouracil with urea and uracil with urea

The two structures have very similar unit cells with only slight differences in some of the dimensions and have the same space group of C2/c (Table 5.8). It might be expected that the introduction of a larger fluorine atom in the 5-fluorouracil molecule instead of the hydrogen in the uracil molecule would result in a larger unit cell. The unit cell volume for the 5-fluorouracil urea molecular complex is indeed larger than that of the uracil urea molecular complex ($1510.22(3) \text{ \AA}^3$ c.f. $1475.37(5) \text{ \AA}^3$), but the expansion is not isotropic. The *a* and *c* axes are longer in the 5-fluorouracil urea molecular complex but the *b* axis is shorter than in the uracil urea molecular complex.

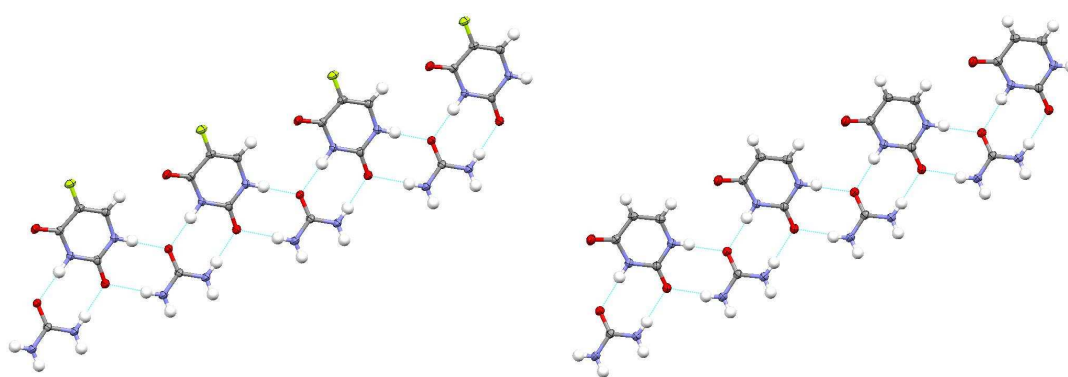


Figure 5.40. Hydrogen bonded chains in left, the 5-fluorouracil and urea molecular complex and right, the uracil and urea molecular complex.

These two structures both contain the same basic chain made up alternating base-pairing between the urea and the uracil/5-fluorouracil molecules (Figure 5.40) comprised of pairs of N-H...O hydrogen bonds. The first is between the amine group of the urea and the oxygen atom in-between the two nitrogen atoms of the ring of the uracil/5-fluorouracil molecule. The second hydrogen bond is from one of the hydrogen atoms on a ring nitrogen atom of the uracil to the oxygen of the urea. The hydrogen bond distances are very similar in both molecular complexes (Table 5.6).

Table 5.6. Comparison of the distances define the base pair motif in the uracil with urea and the 5-fluorouracil and urea molecular complexes.

Hydrogen Bond	5-Fluorouracil and Urea	Uracil and Urea
O---H-N	2.779(3) Å	2.789(4) Å
N-H---O	2.865(3) Å	2.884(4) Å

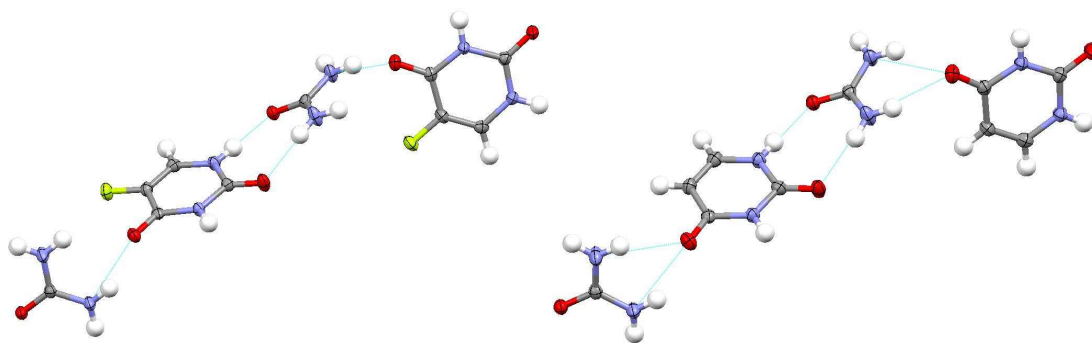


Figure 5.41. Hydrogen bonding between chains in the molecular complexes of left, 5-fluorouracil and urea and right, uracil and urea.

There is a slight difference in these two structures with regards to how the urea orientates itself with respect to the original chain, influenced by the fluorine atom on the 5-fluorouracil molecule. In the uracil urea molecular complex, the urea molecule forms a DDHHA bifurcated hydrogen bond to a carbonyl oxygen on a uracil molecule in a neighbouring chain (Figure 5.41, right). This bifurcated hydrogen bond is also seen in the 5-fluorouracil urea molecular complex but the minor component is significantly longer than that seen in the uracil structure (Table 5.7). This is a consequence of this hydrogen atom forming a second weak N-H...F hydrogen bond with the fluorine atom on the same molecule with an N...F distance of 3.218(2) Å (Figure 5.42). This changes the orientation of the urea molecule relative to the 5-fluorouracil molecule. The significance of this subtle difference is only apparent as the structure is extended (Figure 5.43). When the structures are overlaid with one another (Figure 5.43) the slippage of the molecules relative to one another over longer distances can be clearly seen. This is a consequence of the additional interactions formed involving the fluorine atom (Figure 5.44).

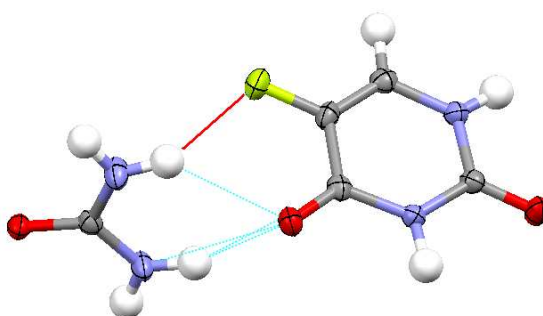


Figure 5.42. Fluorine interaction (red dotted line) and bifurcated hydrogen bond (blue dotted line) in the molecular complex of 5-fluorouracil and urea.

Table 5.7. Comparison of the distances in the bifurcated hydrogen bonds in the uracil with urea and the 5-fluorouracil and urea molecular complexes.

	O---N-H1	O---N-H2
5-Fluorouracil and Urea	3.011(3) Å	3.134(3) Å
Uracil and Urea	3.002(3) Å	2.926(3) Å

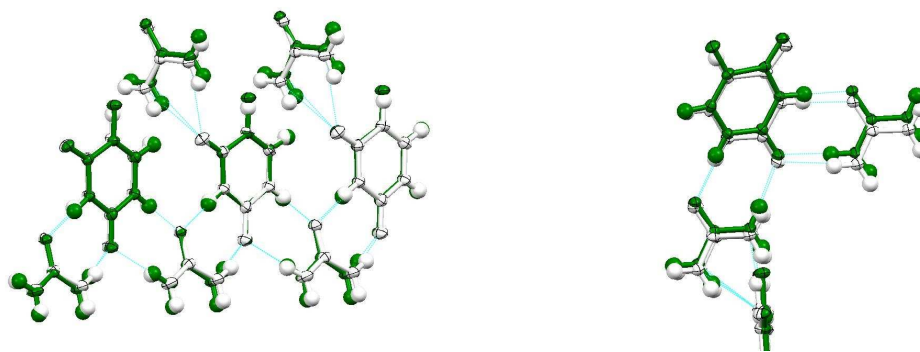


Figure 5.43. Structure overlay of the molecular complexes of 5-fluorouracil and urea (green) and uracil and urea (white) showing divergence over longer length scales.

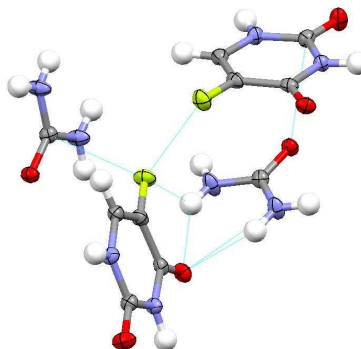


Figure 5.44. The interactions involving the fluorine atom in the molecular complex of 5-fluorouracil and urea.

The fluorine is involved in a second interaction, a fluorine-fluorine interaction between two 5-fluorouracil molecules in different chains. This interaction has an F...F distance of 2.666(3) Å and is shorter than the sum of the van der Waals radii of two fluorine atoms of 2.94 Å. It is the combination of these two fluorine interactions that causes the two related structures (uracil/urea and 5-fluorouracil/urea) to differ marginally.

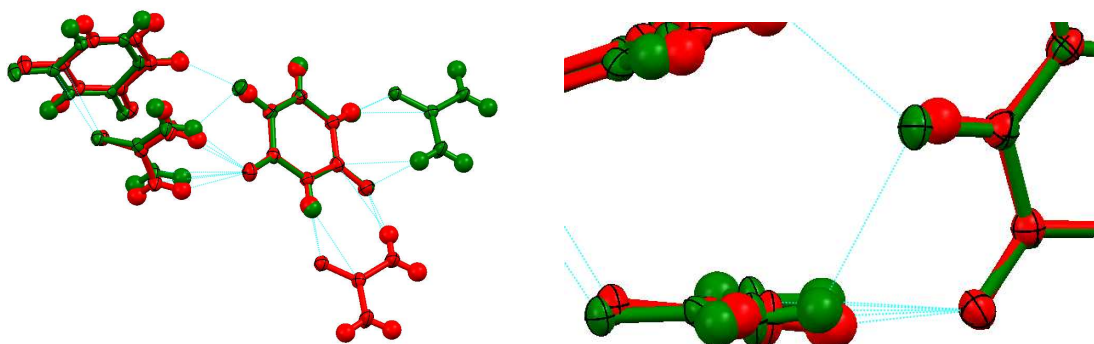


Figure 5.45. Overlaid structures of uracil and urea (red) and the 5-fluorouracil and urea (green) showing subtle interactions due to the fluorine atom.

As can be seen from Figure 5.45, the structures start to differentiate at a very early stage in the expansion of the structure. The fluorine-fluorine interaction in the green structure causes the centroids of the rings of the 5-fluorouracil molecules to move apart. The interaction between the fluorine and the amine group in the green structure changes the orientation of the urea molecule slightly in comparison to the non-fluorinated structure (red). The urea molecule is tilted slightly with the hydrogen involved in a weak interaction

with the fluorine being pulled closer to the fluorine atom, however, with the oxygen atom of the urea being held in place by moderate hydrogen bonds the urea molecule is slightly tilted. This becomes exaggerated as the two structures are expanded and the combination of the two interactions combines to move the two structures out of sync.

5.6.2. The Effect of Sulphur on the Molecular Complexes of Urea and Thiourea with 5-Fluorouracil

The two molecular complexes of 5-fluorouracil and urea, and 5-fluorouracil with thiourea both form in a 1:1 ratio. In this case, the primary hydrogen bonding units are quite different. The urea complex contains hetero- pseudo-base paired motifs between urea and 5-fluorouracil, whilst the complex containing thiourea contains a homo-base paired motif between 5-fluorouracil molecules (Figure 5.46).

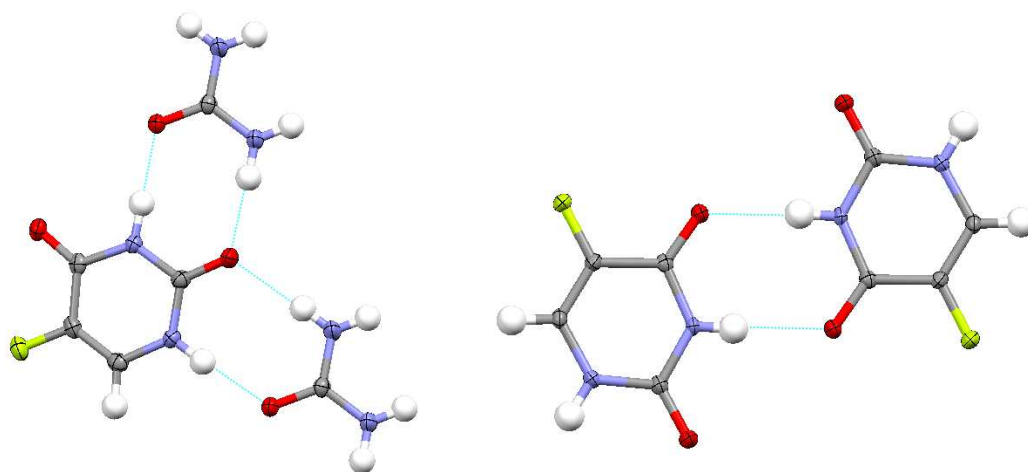


Figure 5.46. Left hetero- pseudo-base pairing in the 5-fluorouracil and urea molecular complex and right, homo-base pairing in the 5-fluorouracil and thiourea molecular complex.

Hydrogen bonded chains of 5-fluorouracil and urea/thiourea are formed in both molecular complexes. However, in the urea complex, these chains are hetero pseudo-base paired units with alternating urea and 5-fluorouracil molecules (Figure 5.47, left). In the thiourea complex, these chains are alternating pairs of 5-fluorouracil and thiourea molecules, where the 5-fluorouracil molecules form homo-base pairs with themselves (Figure 5.47, right).

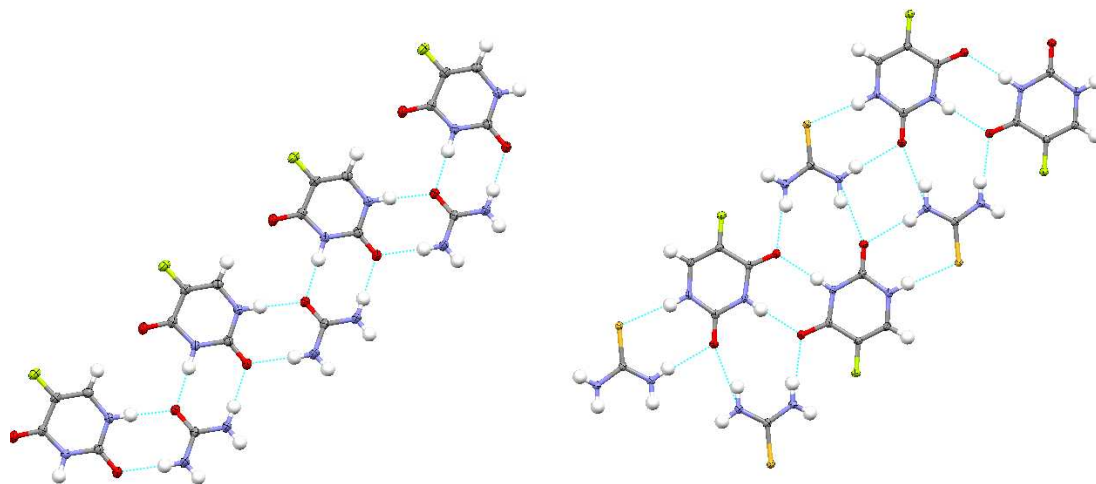


Figure 5.47. Left, extended chain in the 5-fluorouracil and urea molecular complex and right, extended chain in the 5-fluorouracil and thiourea molecular complex.

The thiourea molecule preferentially hydrogen bonds to one area of the 5-fluorouracil molecule and not the other side. The pseudo-base pair motif it forms is with the oxygen on the 5-fluorouracil that is opposite to the fluorine and the hydrogen on the heteroatom opposite the other oxygen atom. This is not the same as the urea structure as here there is a urea molecule bonded to each side of the 5-fluorouracil molecule (Figure 5.48).

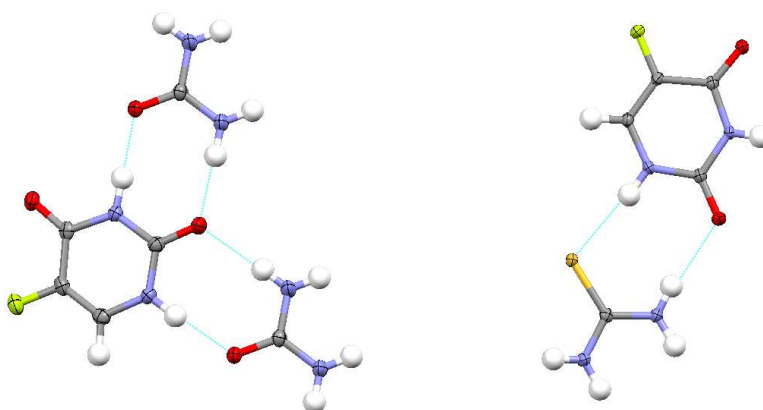


Figure 5.48. Hetero-base pair hydrogen bonding motifs, left on both sides of 5-fluorouracil molecule in the 5-fluorouracil and urea molecular complex and right, only on one side of the 5-fluorouracil molecule in the 5-fluorouracil and thiourea molecular complex.

Due to the changes in the stronger hydrogen bonded motifs in the molecular complexes, the fluorine interactions have been disrupted. In the urea structure the fluorine forms two fluorine interactions, one fluorine-fluorine interaction and one weak hydrogen bond with two separate molecules. However, these interactions are not present in the thiourea molecular complex and the fluorine is only involved in one interaction between the fluorine atom and the amine group of the thiourea. This change in importance of the fluorine atom shows that the presence of the sulphur has caused the fluorine to play a less important role in the extended structure of the thiourea 5-fluorouracil molecular complex. It would therefore appear that the sulphur interactions form preferentially compared to the fluorine interactions in these structures.

5.6.3 Cytosine Based Molecular Complexes

The two molecular complexes including 5-fluorocytosine do show some similar structural motifs despite forming in different ratios. Both form wholly 5-fluorocytosine chains formed by alternating type A and type B base pairs. This is a common trend in the cytosine molecular complexes which do not exhibit hydrogen transfer. In the urea molecular complex there is only one unique chain formed whilst in the thiourea molecular complex, there are two independent 5-fluorocytosine chains formed. The difference here comes down to the difference in the ratio of crystallisation. In the urea molecular complex, there are two independent 5-fluorocytosine molecules and these combine alternately to form each chain. However, in the thiourea molecular complex, there are three independent 5-fluorocytosine molecules. Two of these combine alternately to make up a chain similar to that found in the urea 5-fluorocytosine molecular complex. However, in addition, there is a second independent chain comprised of the final independent 5-fluorocytosine molecule.

One major difference between these structures is the role of the co-molecule. In the molecular complex with thiourea hydrate, the thiourea molecule acts as a linker between separate chains in combination with the water molecule as discussed in Section 5.4 (Figure 5.48, left). This thiourea molecule does not form an extended chain. The molecular complex containing urea forms an extended chain with other urea molecules (Figure 5.48, right). These urea molecules are linked via hydrogen-bonded base paired rings. This is a major difference present in the two structures of the co-molecule.

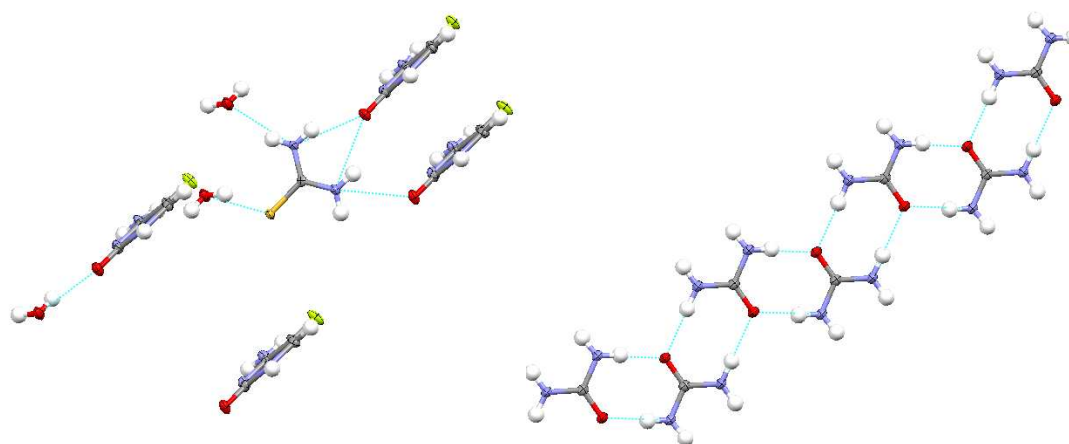


Figure 5.48. Left, thiourea acting as a linker in the 5-fluorocytosine and thiourea hydrate molecular complex and right urea chain built up in the 5-fluorocytosine and urea molecular complex.

There is a similarity in the fluorine interactions in both of these molecular complexes. Weak hydrogen bonds are formed in both which form a unit which resembles a base paired bonding motif comprised of C-H...F hydrogen bonds (Figure 5.49). This is possible due to

the orientation of the fluorine atom relative to the chains in each of the molecular complexes.

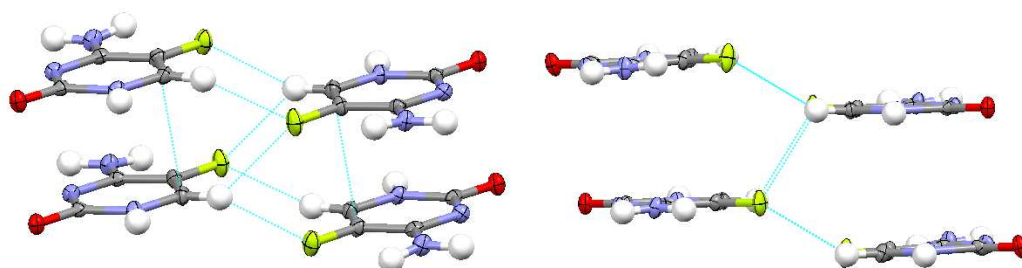


Figure 5.48. Fluorine hydrogen bonds in the molecular complexes of left, 5-fluorocytosine and thiourea hydrate and right, 5-fluorocytosine and urea.

5.7. Conclusions

Solubility issues have restricted the number of molecular complexes that could be obtained in the series of molecular complexes formed from uracil and cytosine with urea and thiourea. However, some common themes could still be identified, all of which involve the formation of base pairs and pseudo base-pair motifs. The addition of the fluorine does not interrupt this base-pairing and only acts as a stabilising factor and a link between different areas of the structure such as linking of separate chains.

The fluorine can have a big impact or play a more subtle role in a structure; the molecular complex of 5-fluorouracil with urea is an example of the latter, when compared to the molecular complex of uracil and urea. The fluorine atom helps to slowly move the two structures out of sync as the basic building block (uracil/5-fluorouracil with urea chains) remains predominantly the same. The fluorine is involved in two interactions and it is the presence of these that cause the slight differentiation between the two structures.

There are obvious differences between the structures that contain thiourea and urea. These molecules provide different bonding opportunities and as a result the presence of either causes differences between the two structures. In the structures obtained it appears that the sulphur (present in the thiourea and 5-fluorocytosine complex) interactions are preferably formed when compared to the fluorine interactions (present in the urea and 5-fluorocytosine complex). However, the types of hydrogen binding interactions in this pair of structures is still similar.

Compound	5-fluorouracil and urea	Uracil and Urea	5-Fluorouracil and Thiourea	5-fluorocytosine and thiourea hydrate	5-fluorocytosine and urea
Formula	C5 H7 F1 N4 O3	C5 H8 N4 O3	'C5 H7 F N4 S O2'	C13 H18 F3 N11 O4 S1	C9 H12 F2 N8 O3
Crystallisation Conditions	Methanol, Room Temperature	Methanol, Room Temperature	Methanol, Room Temperature	Methanol, Room Temperature	Methanol, Room Temperature
Molecular weight / gmol ⁻¹	318.27	172.15	206.2	205.22	318.27
Temperature (K)	100 K	100 K	100 K	100 K	100 K
Space Group	C2/c	C2/c	P-1	P-1	P-1
<i>a</i> (Å)	9.4046(2)	9.2546(2)	6.17590(40)	10.2215(4)	6.8980(4)
<i>b</i> (Å)	10.4073(3)	10.4378(2)	7.11960(40)	10.5680(4)	10.1538(6)
<i>c</i> (Å)	15.6782(4)	15.4590(3)	10.68410(59)	10.9209(4)	10.6856(6)
α (°)	90	90	101.9653(31)	75.767(2)	106.257(2)
β (°)	99.934(1)	99.853(1)	98.2481(32)	63.717(2)	102.730(3)
γ (°)	90	90	114.2782(33)	69.380(2)	109.447(3)
Volume (Å ³)	1511.52(3)	1479.30(5)	404.916(44)	984.26(1)	635.57(6)
<i>Z</i>	8	8	2	2	2
θ range (°)	2.6-33.6	2.7-27.485	1-27.485	2.1-26.7	1-27.485
Completeness	99.90%	99.90%	99.20%	97.60%	98.10%
Reflections Collected	22949	11473	10325	21035	15053
Independent	2817	1684	1831	4067	2888
Refln (obs.I>2 σ (I))	1432	1309	1672	3188	1432
<i>R</i> _{int}	0.0692	0.0692	0.0692	0.2283	0.0765
Parameters	146	141	146	361	395
GooF on F ²	1.096	1.139	1.054	1.045	1.0690
<i>R</i> ₁ (Observed)	0.039	0.066	0.027	0.062	0.0691
<i>R</i> ₁ (all)	0.044	0.083	0.032	0.081	0.1691
<i>wR</i> ₂ (all)	0.114	0.230	0.065	0.165	0.1897

Table 5.8. Crystallographic data for the molecular complexes reported in Chapter 5.

Chapter 6 Molecular Complexes of Disubstituted Benzoic acids with Cytosine, Uracil and their Fluorinated Derivatives

6.1. Polymorphic 1:1 Molecular Complexes of 5-Fluorocytosinium 2,6-Dihydroxybenzoate

6.1.1 1:1 Molecular Complex of 5-Fluorocytosinium 2,6-Dihydroxybenzoate Form I

5-Fluorocytosine and 2,6-dihydroxybenzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.3mm x 0.3mm x 0.3mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 6.5.

This molecular complex forms in a 1:1 ratio and there is 100% hydrogen transfer from the 2,6-dihydroxybenzoic acid molecule to the 5-fluorocytosine. The primary bonding motif adopted in this structure is therefore a hydrogen-bonded $R_2^2(8)$ ring motif between the two components (Figure 6.1). The hydrogen-bonded ring is comprised of two moderate strength N-H...O hydrogen bonds. The first hydrogen bond is from one of the hydrogen atoms of the amine group of the 5-fluorocytosinium to one of the oxygen atoms of the carboxylate group of the 2,6-dihydroxybenzoate. This hydrogen bond has an N...O distance of 2.7994(17)Å. The second hydrogen bond is from the hydrogen on the protonated nitrogen of the cytosine to the other oxygen atom of the carboxylate group. This bond has an N...O distance of 2.6534(17)Å. These hydrogen bonds are both charge assisted and of moderate strength.

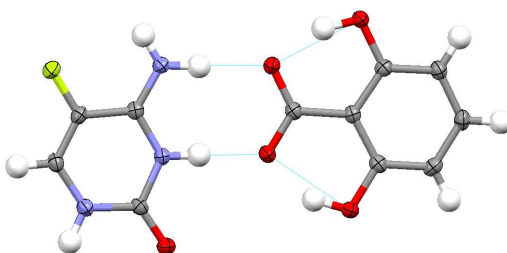


Figure 6.1. Heterodimer formed through a hydrogen-bonded ring between 5-fluorocytosinium and 2,6-dihydroxybenzoate molecules in Form I.

Each heterodimer is hydrogen bonded to four other dimers (Figure 6.2). There are two unique N-H...O hydrogen bonds: one is from a hydrogen of the amine group of a 5-fluorocytosinium molecule to the oxygen of one of the hydroxy groups of the 2,6-dihydroxybenzoate and has an N...O distance of 2.8181(18)Å; the second is from a hydrogen on a ring nitrogen of a 5-fluorocytosinium molecule to the oxygen of the second hydroxy group of a 2,6-dihydroxybenzoate molecule and has an N...O distance of

2.8502(18)Å. These dimers are orientated at $\sim 45^\circ$ to one another and generate corrugated layers.

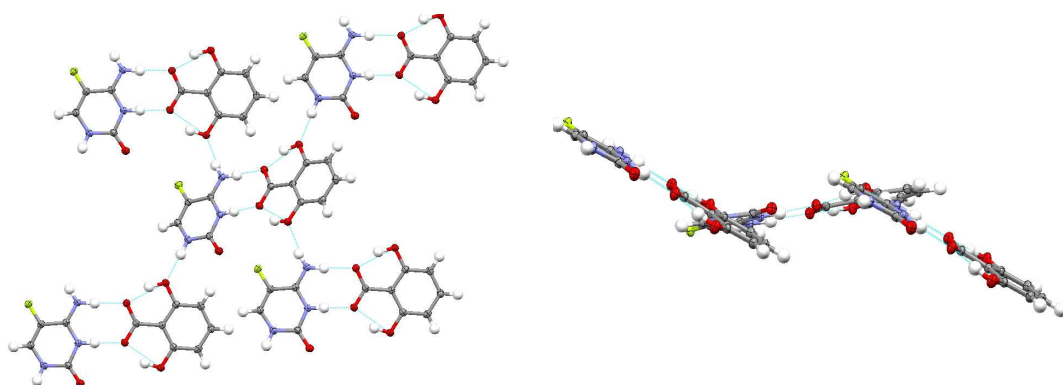


Figure 6.2. Hydrogen bonded heterodimers assembling into corrugated layers in the molecular complex of 5-fluorocytosinium 2,6-dihydroxybenzoate Form I.

Separate, parallel heterodimers are linked via a bonding motif that resembles a stepped base pair (Figure 6.3). This motif is made up of two weak C-H \cdots F hydrogen bonds involving the fluorine of one fluorocytosinium molecule and the hydrogen on an aromatic carbon of the second fluorocytosinium molecule. This bond is reciprocated to complete the base-pair motif. Both of these bonds are of weak strength and have C \cdots F distances of 3.514(2) Å.

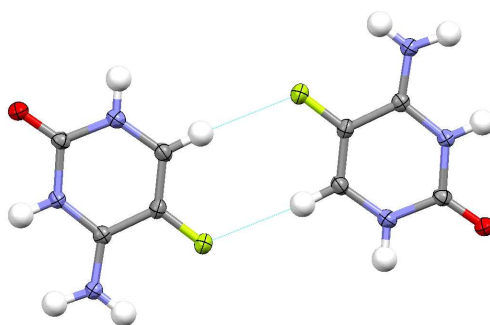


Figure 6.3. Base-pair style motif involving fluorine in the molecular complex of 5-fluorocytosinium 2,6-dihydroxybenzoate Form I.

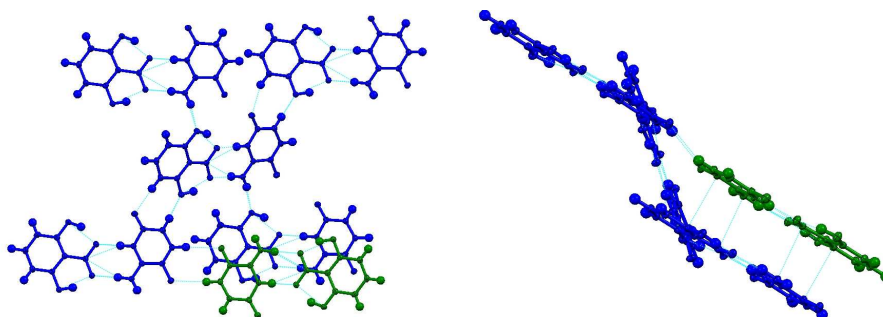


Figure 6.4. Fluorine interactions linking base pairs in parallel stepped layers in the molecular complex of 5-fluorocytosinium 2,6-dihydroxybenzoate Form I.

The charged fluorocytosinium molecules are stacked on top of the 2,6-dihydroxybenzoate molecules with an approximate stacking distance of 3.276 Å. This type of stacking is

common in molecular complexes containing charged species. The global packing illustrates that the structure is organised with the separate regions of fluorocytosinium (green) and 2,6-dihydroxybenzoate (blue) (Figure 6.5). The alternate layering of 5-fluorocytosinium and 2,6-dihydroxybenzoic acid molecules is along the *b*-axis.

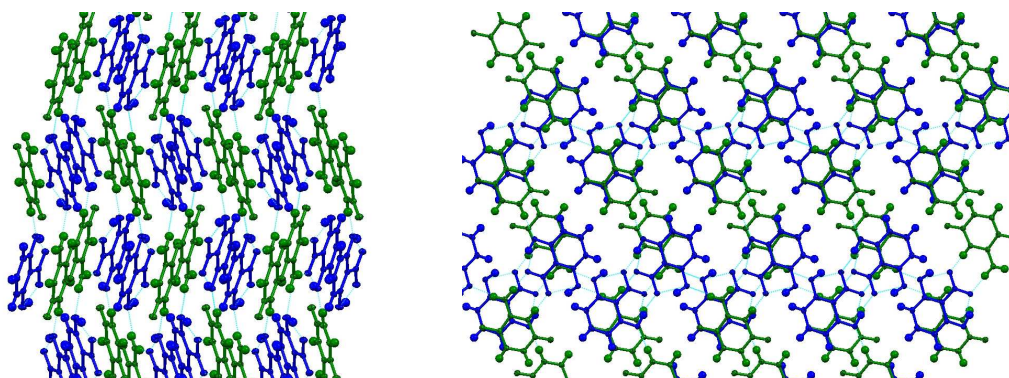


Figure 6.5. Stacking interactions in the molecular complex of 5-fluorocytosinium 2,6-dihydroxybenzoate Form I (5-fluorocytosinium in green and 2,6-dihydroxybenzoate in blue).

6.1.2 1:1 Molecular Complex of 5-Fluorocytosinium 2,6-Dihydroxybenzoate Form II

5-fluorocytosine and 2,6-dihydroxybenzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.3mm x 0.3mm x 0.3mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 6.5.

The polymorph, Form II, shows some structural similarities to Form I. The molecular complex forms in a 1:1 ratio with 100% hydrogen transfer from the carboxyl group of a 2,6-dihydroxybenzoic acid to a ring nitrogen of a 5-fluorocytosine to create 5-fluorocytosinium and a 2,6-dihydroxybenzoate molecules. The heterodimer, held together by a $R_2^2(8)$ hydrogen bonded ring, is again formed.

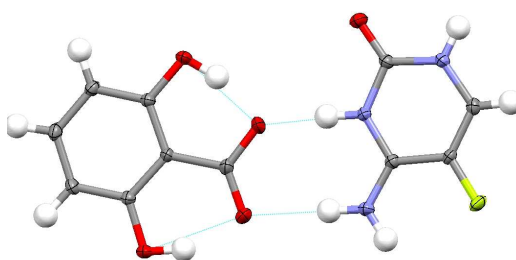


Figure 6.6. Hydrogen-bonded ring formed between 5-fluorocytosinium and 2,6-dihydroxybenzoate in the Form II molecular complex.

There are two N-H...O hydrogen bonds in this hydrogen bonded ring. The hydrogen bond donor of the first is the amine group and the second is a hydrogen on a ring nitrogen both to the carboxylate group of the 2,6-dihydroxybenzoate molecule. The hydrogen bonds are of moderate strength and have N...O distances of 3.0426(14) Å and 2.6223(13) Å, respectively. The first is significantly longer than that observed in Form 1 and the second is slightly shorter. The relative orientation of the 5-fluorocytosinium and 2,6-dihydroxybenzoate molecules is therefore different in the two forms (Figure 6.7).

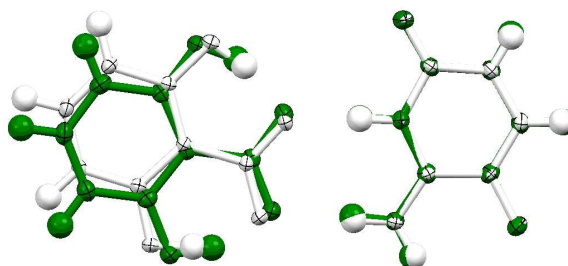


Figure 6.7. An overlay of the heterodimers of the Form I (green) and Form II (white) molecular complexes of 5-fluorocytosinium 2,6-dihydroxybenzoate.

As seen in Form I, each heterodimer is hydrogen bonded to four other dimers (Figure 6.8, left). Once again there are two unique N-H...O hydrogen bonds: one is from a hydrogen of the amine group of a 5-fluorocytosinium molecule to the oxygen of one of the hydroxyl groups of the 2,6-dihydroxybenzoate and has an N...O distance of 2.986(2) Å; the second is from a hydrogen on a ring nitrogen of a 5-fluorocytosinium molecule to the oxygen of the second hydroxyl group of a 2,6-dihydroxybenzoate molecule and has an N...O distance of 2.763(2) Å. These dimers form almost planar layers.

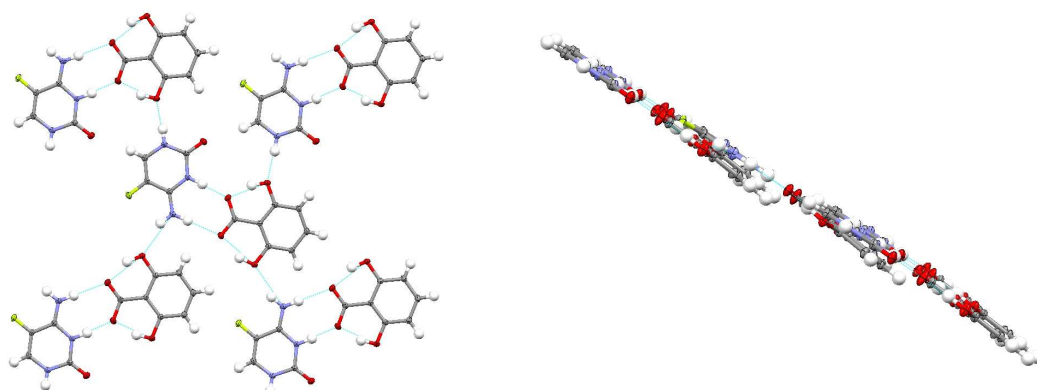


Figure 6.8. Hydrogen bonded heterodimers assembling into planar layers in the molecular complex of 5-fluorocytosinium 2,6-dihydroxybenzoate Form II

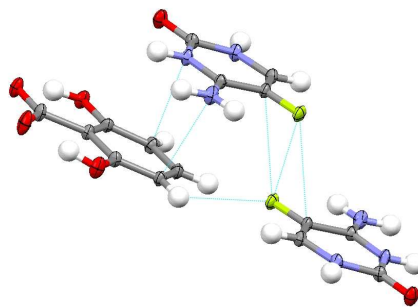


Figure 6.9. Fluorine interactions in the molecular complex of 5-fluorocytosinium 2,6-dihydroxybenzoate Form II.

This structure is more layered than the corrugated structure of Form I. This change in orientation of adjacent heterodimers results in a change in the fluorine interactions that are formed. The fluorine atom in Form II is involved in two interactions (Figure 6.9). The first is between two fluorine atoms in parallel layers. The distance between the two fluorine atoms is 2.642 (2) Å, which is significantly shorter than the sum of the van der Waals radii for two fluorine atoms (2.94 Å). The second interaction is a weak C-H...F hydrogen bond between an aromatic hydrogen atom of a neighboring 2,6-dihydroxybenzoate molecule with a C...F distance of 3.210(2) Å. There is no base pair comprised of C-H...F hydrogen bonds in Form II.

Again, the charged 5-fluorocytosinium molecules stack on top of only 2,6-dihydroxybenzoate molecules and no 5-fluorocytosinium molecules are stacked directly on other charged 5-fluorocytosinium molecules (Figure 6.10). The approximate stacking distance is 3.200 Å.

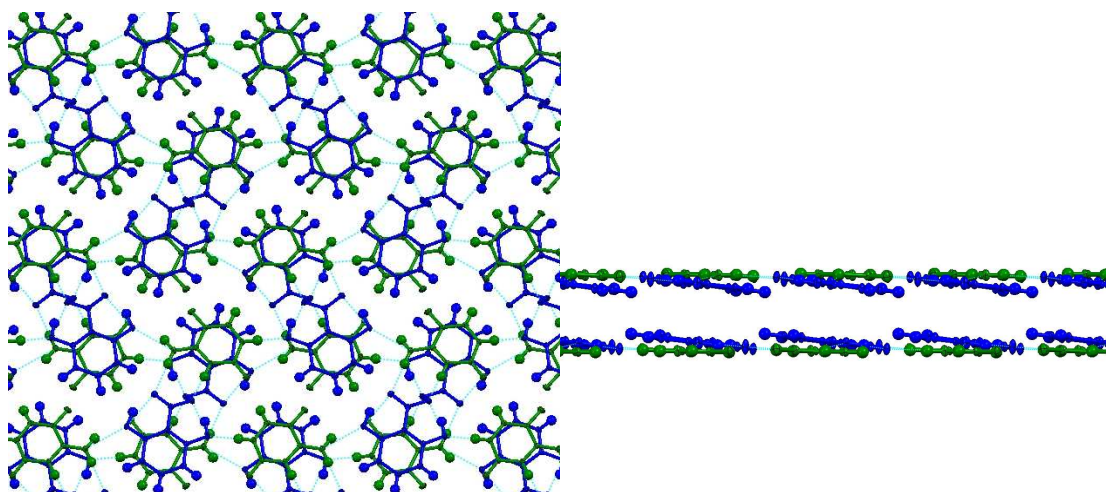


Figure 6.10. Stacking interactions in the molecular complex of 5-fluorocytosinium 2,6-dihydroxybenzoate Form II.

6.1.3 Structural Comparison of the Two Polymorphs of 5-Fluorocytosinium 2,6-Dihydroxybenzoate

The major difference between forms I and II is the construction of the layers. Form I has a corrugated structure that contains a twisting of adjacent heterodimers whilst Form II forms into a much more planar layered structure (Figure 6.10). There is a moderate strength C-H...F hydrogen bond between the oxygen atom of the 5-fluorocytosinium molecule and one an aromatic hydrogen atom of a nearby 5-fluorocytosine molecule in Form II which is not present in Form I.

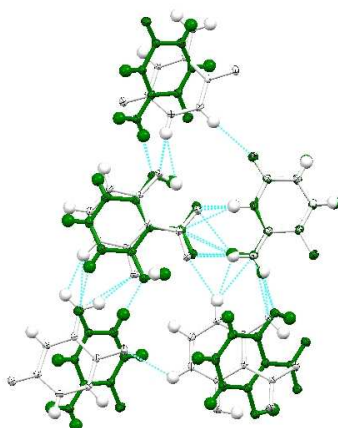


Figure 6.11. An overlay of the molecular complexes of Form I (green) and Form II (white) of 5-fluorocytosinium 2,6-dihydroxybenzoate showing the different construction of the layers.

There is an additional change in the relative orientation of neighbouring heterodimers. In Form II, the 5-fluorocytosinium molecules on both sides of the heterodimer are flipped relative to those in Form I (Figure 6.11). This enables the formation of the C-H...O hydrogen bond on one side, and results in the interaction with the 2,6-dihydroxybenzoate molecule on the other side of the dimer to be through the amine group rather than the protonated ring nitrogen in Form I. The fluorine interactions in the two polymorphs are thus also significantly different. However, both show the same stacking interactions.

6.2 2:1:1 Molecular Complex of Cytosine and 2,6-Dihydroxybenzoic acid Hydrate

Cytosine and 2,6-dihydroxybenzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.3mm x 0.3mm x 0.3mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 6.5.

The molecular complex forms in a 2:1:1 ratio of cytosine to 2,6-dihydroxybenzoic acid to water. As a result, there is one neutral cytosine, one protonated cytosinium, a 2,6-dihydroxybenzoate and a water molecule in the asymmetric unit. In common with the other structures reported in this thesis, the presence of both protonated and neutral cytosine molecules results in the pseudo-Watson-Crick hydrogen bonding motif (Figure 6.12). This forms in preference to the heterodimer formed in the polymorphs of 5-fluorocytosinium with 2,6-dihydroxybenzoate. The pseudo-Watson-Crick bonding motif is made up of three hydrogen bonds, two of which are N-H...O hydrogen bonds with the remaining bond being an N...N hydrogen bond. The hydrogen bond coming from the amine group of the cytosinium involving the oxygen atom of the cytosine has an N...O distance of 2.7654(10) Å. The next bond is from the protonated heteroatom of the cytosinium ring to the non-protonated heteroatom of the cytosine ring. This bond has an N...N distance of 2.8188(10) Å and is of moderate strength. The final hydrogen bond is from the amine group of the cytosine to the oxygen atom of the cytosinium. This bond is of moderate strength and has an N...O distance of 2.8900(11) Å.

Chains of cytosine molecules are formed through two different hydrogen bonding motifs, the pseudo-Watson-Crick, and the type B pseudo Hoogsteen formed by two N-H...O hydrogen bonds (Figure 6.12). with N...O distances of 2.7958(11)Å (from the cytosinium) and 2.7633(11)Å (from the cytosine).

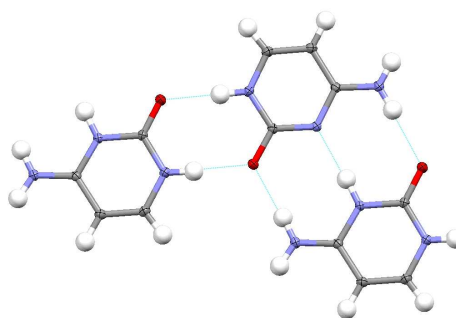


Figure 6.12, Chains of alternating Type B and pseudo-Watson-Crick base pairs in the hydrated molecular complex of cytosine and 2,6-dihydroxybenzoic acid.

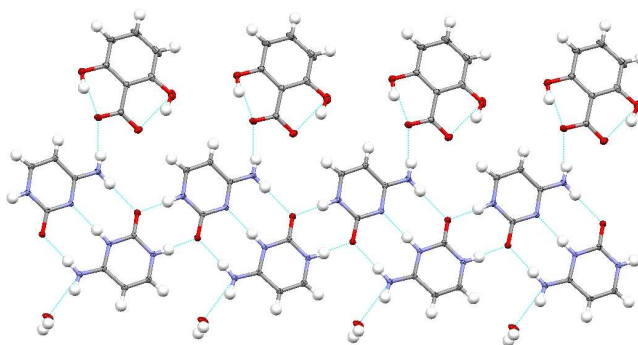


Figure 6.13. The hydrogen bonding to the cytosine chains in the molecular complex of cytosine and 2,6-dihydroxybenzoic acid hydrate. The benzoic acid molecules lie on one side of the chain and the water molecules on the other side.

The 2,6-dihydroxybenzoate molecules are hydrogen bonded to the chain via a single hydrogen bond from the amine group of a cytosine to one of the carboxylate group oxygen atoms (Figure 6.13). This bond is of moderate strength and has an N...O distance of 2.9101(11) Å. The 2,6-dihydroxybenzoate molecule lies in the same plane as the cytosine chain. On the other side of the chain, the water molecule is tied to the chain via single hydrogen bond from the amine group of the cytosinium to the oxygen of the water molecule (Figure 6.13). This bond is of moderate strength and has an N...O distance of 2.8035(11) Å. The oxygen atom of the water molecule lies slightly out of the plane of the chain and the molecule lies almost perpendicular to the chain forming a connecting role between different chains (Figure 6.14).

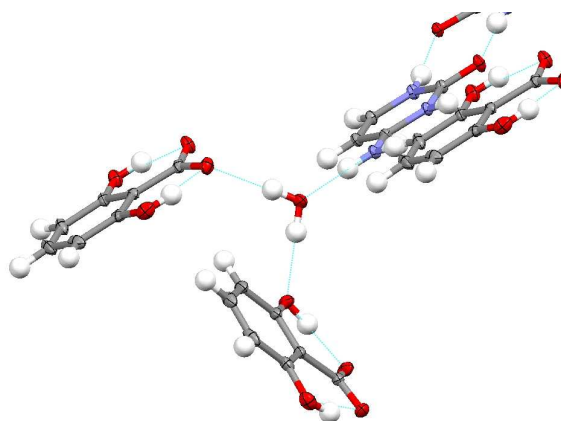


Figure 6.14. Hydrogen bonding involving the water molecule in the cytosine and 2,6-dihydroxybenzoic acid hydrate molecular complex.

The water molecule acts as a hydrogen bond donor to a 2,6-dihydroxybenzoate molecule in two separate chains. The first of these hydrogen bonds is to a oxygen atom on the carboxylate group of a 2,6-dihydroxybenzoate molecule in a parallel chain. This hydrogen bond has an O...O distance of 2.8960(11) Å and is of moderate strength. The second water hydrogen forms a hydrogen bond to the hydroxyl group of a 2,6-dihydroxybenzoate molecule in an almost perpendicular chain. This bond is of moderate strength and has an O...O distance of 2.8417(11) Å.

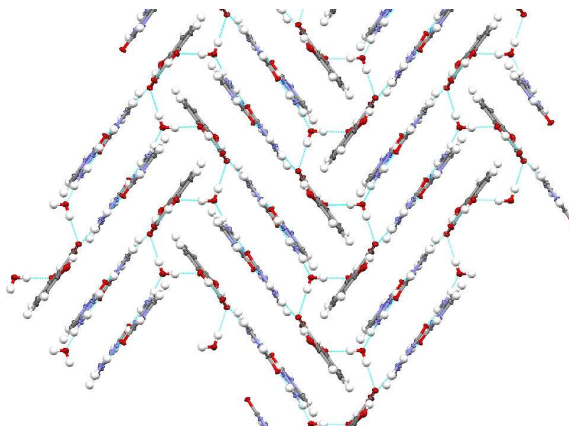


Figure 6.15. Herring-bone arrangement along the *a*-axis in the molecular complex of cytosine and 2,6-dihydroxybenzoic acid hydrate.

The combination of the different orientation of the chains produces a herring-bone shape to the structure when viewed along the *a*-axis with regions of base stacking (Figure 6.15). Cytosine molecules stack on cytosinium molecules in a parallel cytosine chain and the 2,6-dihydroxybenzoate molecules sandwich these (Figure 6.18). The stacking distance found between the cytosine layers is approximately 3.286 Å and the two between cytosine/cytosinium and 2,6-dihydroxybenzoate molecules are approximately 3.324 Å and 3.300 Å.

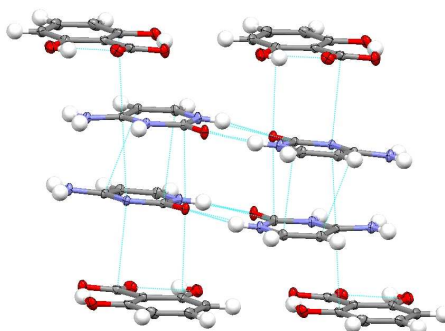


Figure 6.16. Stacking interactions in the molecular complex of cytosine and 2,6-dihydroxybenzoic acid hydrate.

6.3. Polymorphs of the 1:1:1 Molecular Complex of Cytosinium 2,4-Dihydroxybenzoate Hydrate

6.3.1 1:1:1 Molecular Complex of Cytosinium 2,4-Dihydroxybenzoate Hydrate Form I

Cytosine and 2,4-dihydroxybenzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.3mm x 0.3mm x 0.3mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 6.5.

The molecular complex forms in a 1:1:1 ratio of cytosine to 2,4-dihydroxybenzoic acid to water. In this structure there is 100% hydrogen transfer from the 2,4-dihydroxybenzoic acid to the cytosine molecule.

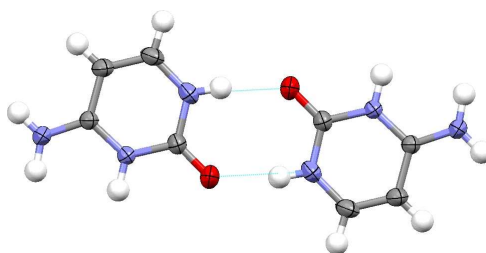


Figure 6.17. Type B base-pairing in the cytosinium 2,4-dihydroxybenzoate hydrate Form I molecular complex.

There is one homo-base-pairing motif in this structure, between the oxygen of one cytosine molecule and a protonated ring nitrogen of a neighbouring cytosine molecule (Figure 6.17). These N-H...O hydrogen bonds are of moderate strength and have N...O distances of 2.9182(15) Å. The base-paired unit is not planar, adopting a stepped configuration with an inversion centre located in the middle of the hydrogen bonded ring. Chains of cytosinium do not form in this case, instead forming a hydrogen bonded ring with a 2,4-dihydroxybenzoate molecule (Figure 6.18) in the position where the type A base pairing would be present, creating a four molecule unit.

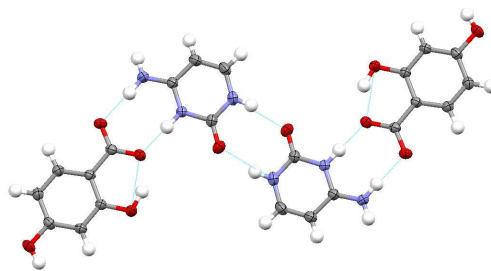


Figure 6.18. Termination of cytosine chains via hydrogen-bonded ring formation creating a four molecule unit in the cytosinium 2,4-dihydroxybenzoate hydrate Form I molecular complex.

The hydrogen bonded ring involves two N-H...O hydrogen bonds with the cytosinium molecule acting as the hydrogen bond donor in both cases. One of these bonds is from the amine group of the cytosinium and the other from a protonated ring nitrogen of the cytosinium with each of these going to one of the oxygen molecules of the carboxylate group. The protonated ring nitrogen hydrogen bonds to the oxygen that is also involved in an intramolecular hydrogen bond with the hydroxyl group of the 2,4-dihydroxybenzoate molecule. This hydrogen bond is of moderate strength and has an N...O distance of 2.6601(12) Å. The other hydrogen bond involving the amine group has an N...O distance of 2.7856(13) Å and this is also of moderate strength.

There is one independent water molecule which forms three hydrogen bonds acting as a hydrogen bond acceptor for a hydrogen bond from the *para* hydroxyl group of the 2,4-dihydroxybenzoate molecule and as hydrogen bond donor to both the carbonyl oxygen of the cytosinium and to one of the carboxyl oxygen atom on the 2,4-dihydroxybenzoate molecule (Figure 6.19, left). All of the hydrogen bonds involving the water molecule are O-H...O bonds and of moderate strength. The first is from the hydrogen on the water molecule to the oxygen of the cytosinium molecule with an O...O distance of 2.7875(16) Å. The second hydrogen bond is from the hydrogen on a water molecule to the oxygen of the carboxylate group which is not involved in the intramolecular hydrogen bond with a O...O distance of 2.7609(17) Å. The third hydrogen bond involves the hydrogen of the hydroxyl group in the *para* position and the oxygen of a nearby water molecule with an O...O distance of 2.6495(16) Å and this is the strongest of the three hydrogen bonds. Through these hydrogen bonds, the water molecule acts as a bridging unit between units in different layers.

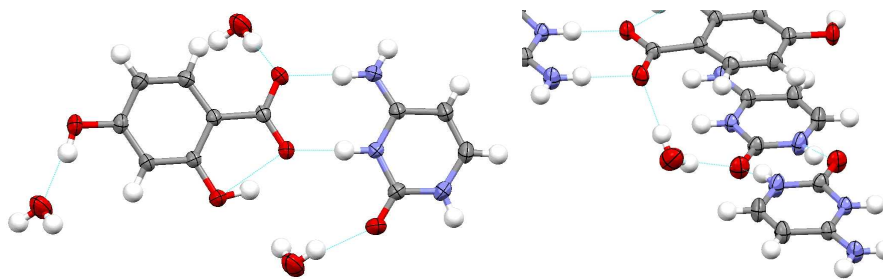


Figure 6.19. Left, the three hydrogen bonds involving water molecules and right, the water molecule connecting the layers in the cytosinium 2,4-dihydroxybenzoate hydrate Form I molecular complex.

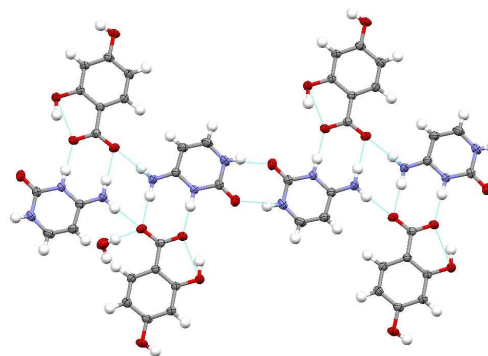


Figure 6.20. Link between four molecule blocks in the cytosinium 2,4-dihydroxybenzoate hydrate molecular complex.

An additional link between the four molecule units is a diamond shaped $R_2^2(8)$ hydrogen bonded ring (Figure 6.20). An inversion centre is located in the middle of this ring and therefore there are two independent hydrogen bonds formed. Both are N-H...O hydrogen bonds from the amine group of the cytosinium molecule to the carboxylate group of the 2,4-dihydroxybenzoate molecule. The first of these hydrogen bonds also forms part of the hydrogen bonded ring of the heterodimer in the four molecule unit. The second connects two stepped layers and is hydrogen bonded to the oxygen of the carboxylate group that is not involved in the intramolecular hydrogen bond. This bond is of moderate strength and has an N...O distance of 2.9482(14) Å.

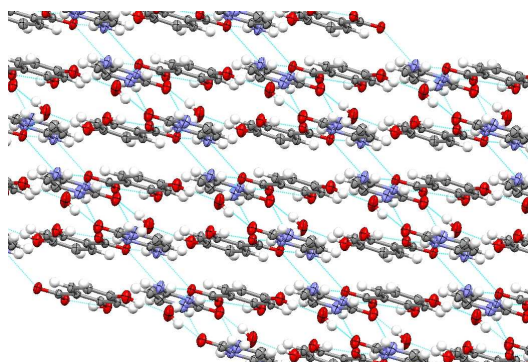


Figure 6.21. The layered structure of the cytosinium 2,4-dihydroxybenzoate hydrate Form I molecular complex.

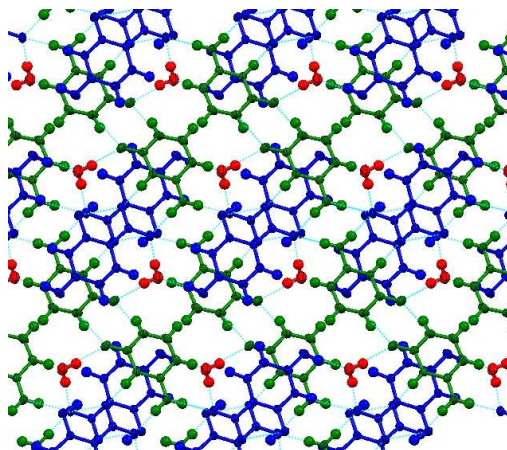


Figure 6.22. Stacking interactions in the molecular complex of cytosinium 2,4-dihydroxybenzoate hydrate Form I. Cytosinium (green), 2,4-dihydroxybenzoate (blue) and water (red).

In addition to the stronger hydrogen bonding interactions which form between layers, there are also base-stacking interactions. Figure 6.22 shows that there are stacking interactions with cytosinium stacked on 2,4-dihydroxybenzoate. This means that there are two base stacking interactions of approximate spacings of 3.333 Å and 3.210 Å.

6.3.2. 1:1:1 Molecular Complex of Cytosinium 2,4-Dihydroxybenzoate Hydrate Form II

Cytosine and 2,4-dihydroxybenzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.3mm x 0.3mm x 0.3mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 6.5.

This molecular complex also crystallises in a 1:1:1 ratio of cytosine to 2,4-dihydroxybenzoic acid to water. It is thus a polymorph of the complex described in Section 6.3.1 and both show structural similarities. The 100% proton transfer again leads to the type B pseudo Hoogsteen base-pairing motif between cytosinium molecules (Figure 6.23). There is an inversion centre in the middle of the hydrogen bonded ring and therefore the two N-H...O hydrogen bonds are equivalent. The hydrogen bonds are of moderate strength and have N...O distances of 2.846(2) Å, shorter than those found in Form I.

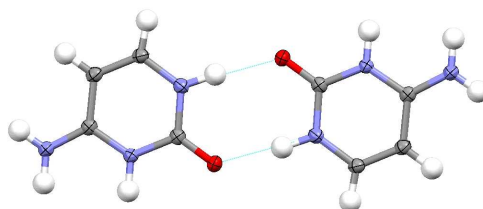


Figure 6.23. Type B base-pairing in the cytosine-2,4-dihydroxybenzoic acid hydrate form II molecular complex.

The cytosinium chain is again terminated by the 2,4-dihydroxybenzoate molecules, which form a hydrogen bonded ring with the cytosinium molecule (Figure 6.24). There are two moderate strength N-H...O hydrogen bonds to the carboxylate group, one from the amine group and the other from a protonated ring nitrogen of the cytosinium molecule. The bond involving the protonated ring nitrogen has an N...O distance of 2.675(2) Å and the hydrogen bond involving the amine group has an N...O distance of 2.763(2) Å. The first of these hydrogen bonds is longer than the equivalent found in Form I and the second is shorter (c.f. 2.6601(12)Å and 2.7856(13)Å, respectively). This corresponds to a deviation from planarity of the four molecule unit in Form II.

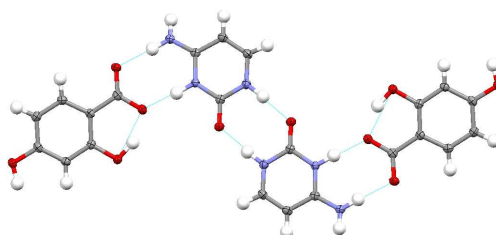


Figure 6.24. Termination of the cytosinium chain via hydrogen-bonded ring formation with the 2,4-dihydroxybenzoate molecules in the cytosinium 2,4-dihydroxybenzoate hydrate Form II molecular complex.

The water molecule forms similar hydrogen bonds to those found in Form 1; there are three unique hydrogen bonds, two as a hydrogen bond donor and one as a hydrogen bond acceptor and these hydrogen bonds interact with the same parts of the four molecule unit (Figure 6.25). However, the orientation of these water molecules is different.

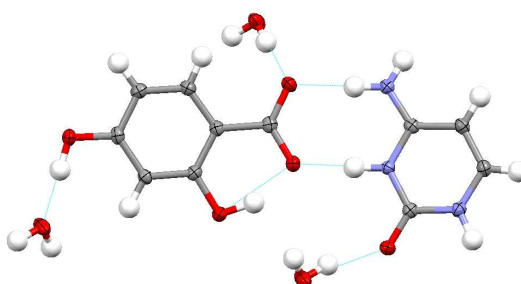


Figure 6.25. Hydrogen bonds involving water molecules in the cytosinium 2,4-dihydroxybenzoate hydrate Form II molecular complex.

All of the hydrogen bonds involving the water molecule are moderate strength O-H...O hydrogen bonds. The first is from the hydrogen on the water molecule to the oxygen of the cytosinium molecule with an O...O distance of 2.7534(19) Å. The second hydrogen bond is from the second hydrogen on the water molecule to the oxygen of the carboxylate group not involved in the intramolecular hydrogen bond with an O...O distance of 2.732(2) Å. The final hydrogen bond is from the hydrogen of the hydroxyl group in the *para* position to the oxygen of a nearby water molecule with an O...O distance of 2.6328(18) Å. These hydrogen bonds are all shorter than those found in Form I.

Figure 6.26 shows the difference between the two forms; the four molecule unit (red) in both structures remains very similar in each. The water molecules also hydrogen bond to the same areas of the molecules and play the same role in the overall structure. The difference from here arises in how the neighbouring heterodimers connect to the four molecule unit.

In Form II, the heterodimer that is linked to the four molecule unit is flipped relative to Form I. In Form I a diamond shape is formed via four hydrogen bonds involving four separate molecules (Figure 6.26, left). In Form II this bonding pattern is not observed and there is no diamond shape in the structure. The link between the next heterodimer and the four molecule building block is through an amine group hydrogen bonding to the hydroxyl group of the benzoic acid and so in this case the heterodimer unit is flipped and so the diamond bonding motif cannot be formed. In Form I there are four molecules involved in the link between the two heterodimers; this is also the case in Form II. This includes a cytosinium, two 2,4-dihydroxybenzoate molecules and a water molecule rather than two cytosinium molecules and two 2,4-dihydroxybenzoate molecules as seen in Form I. The hydrogen bonds here that make up the link are all of moderate strength. The first of these bonds comes from the hydrogen atom on the hydroxyl group in the 4 position of the 2,4-dihydroxybenzoate and involves the oxygen atom of the nearby water molecule with the O...O distance being 2.633(2) Å. The next bond involves a hydrogen atom of the water molecule and one of the oxygen atoms of the carboxylate group of the 2,4-dihydroxybenzoate with this bond having an O...O distance of 2.732(2) Å. The next bond involves the amine group of the cytosinium molecule and the same oxygen atom of the carboxylate group involved in the last hydrogen bond with an O...O distance of 2.763(2) Å. The final hydrogen bond in the link is between the amine group of the cytosinium molecule and the oxygen atom of the hydroxyl group in the 2 position of the 2,4-dihydroxybenzoate molecule with an O...O distance of 2.885(2) Å. The water molecule in Form II shifts its position slightly and this results in the different ring formation.

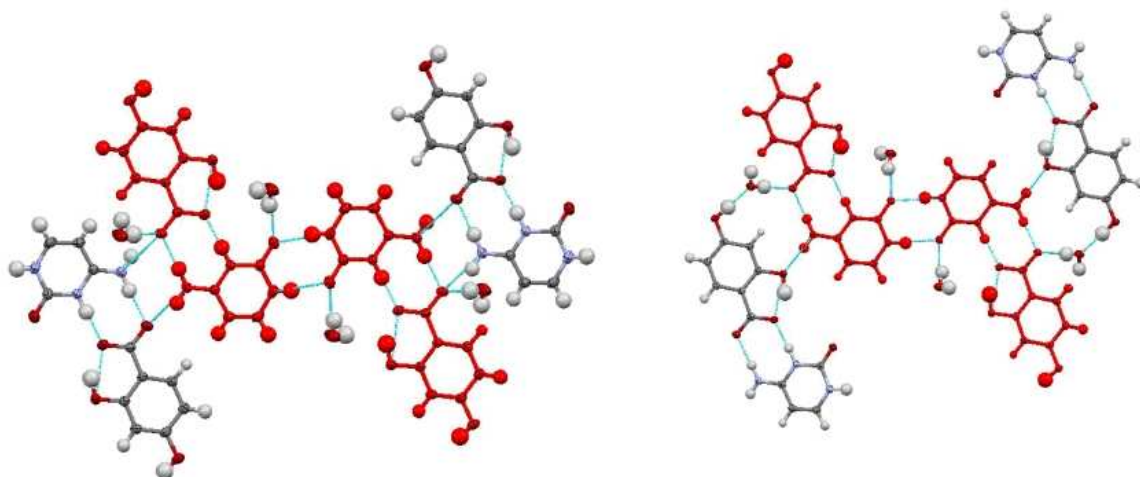


Figure 6.26. Differences in molecules bonded to the four molecule unit (red) in the molecular complexes of cytosinium 2,4-dihydroxybenzoate hydrate Form I (left) and Form II (right).

The final important factor in this structure are the stacking interactions. Again, the cytosinium molecules do not stack on top of other cytosinium molecules and the 2,4-dihydroxybenzoate molecules do not stack on top of each other. The cytosinium molecule stacks on top of the 2,4-dihydroxybenzoate molecules. There are two unique stacking interactions ranging between approximately 3.215-3.348 Å and 3.185-3.341 Å. The ranges are due to the slight twisting of the cytosinium molecule with respect to the layers above and below containing the 2,4-dihydroxybenzoate molecules.

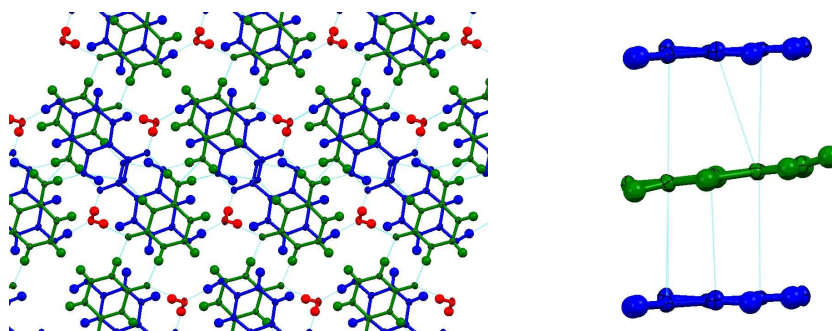


Figure 6.27. Stacking interactions in the molecular complex of cytosinium 2,4-dihydroxybenzoate hydrate Form II.

6.4 2:2:5 Molecular Complex of 5-Fluorocytosinium 3,5-Dihydroxybenzoate Hydrate

5-Fluorocytosine and 3,5-dihydroxybenzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.3mm x 0.3mm x 0.3mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 6.5.

This molecular complex forms in a 2:2:5 5-fluorocytosine to 3,5-dihydroxybenzoic acid to water ratio. Within this, it shows 100% hydrogen transfer from the carboxylate group of the 3,5-dihydroxybenzoate molecules to the unprotonated nitrogen of the 5-fluorocytosinium rings; there are two independent 5-fluorocytosinium and 3,5-dihydroxybenzoate molecules in the asymmetric unit. The pseudo-Watson-Crick base pairing motif is therefore not possible. As might be expected for systems showing this 100% proton transfer, homo-base pairing of the pseudo Hoogsteen type B is found.

This bonding motif is made up of two N-H...O hydrogen bonds between two independent 5-fluorocytosinium molecules and in this case, this produces an approximately planar unit. Both these hydrogen bonds are moderate in strength and are from a protonated ring nitrogen atom to the carbonyl oxygen of the 5-fluorocytosinium but they are inequivalent with N...O distances of 2.8572(16) Å and 2.8615(16) Å (Figure 6.28).

The cytosine chain cannot continue with a type A base pair motif due to the 100% hydrogen transfer. Instead a hydrogen bonded ring is formed between the carboxylate group of the 3,5-dihydroxybenzoate molecules and the 5-fluorocytosinium molecules on both sides of the homo-base pair creating a four molecule unit (Figure 6.28, left). Each hydrogen bonded ring is made up of two N-H...O hydrogen bonds with one coming from the amine group of the 5-fluorocytosinium to one of the oxygen atoms of the carboxylate group and the other comes from a protonated ring nitrogen of the 5-fluorocytosinium to the other oxygen of the carboxylate group. These hydrogen bonds are of moderate strength with N...O distances of 2.7135(15) Å and 2.6999(15) Å, respectively for one side and 2.7619(16) Å and 2.7685(15) Å, respectively for the other side.

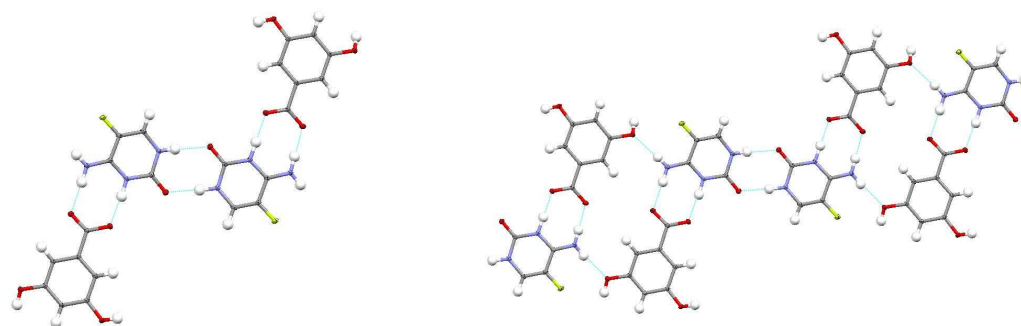


Figure 6.28. Left, four-molecule hydrogen bonded unit and right, two four molecule hydrogen bonded rings linked by type B base pairing in the 5-fluorocytosinium 3,5-dihydroxybenzoate hydrate molecular complex.

Neighbouring four molecule units within the same plane are connected through two moderate strength N-H...O hydrogen bonds that do not involve water molecules. These are from the amine group of a 5-fluorocytosinium molecule to one of the hydroxyl groups on a 3,5-dihydroxybenzoate molecule. This occurs on both sides of the four molecule unit with

N \cdots O distances of 2.8615(16) Å and 2.8860(16) Å, respectively. This forms two four molecule $R_2^4(18)$ hydrogen bonded rings linked to each other via the type B base-pairing (Figure 6.28, right).

There are five independent water molecules in this molecular complex. A water molecule acts as a bridge between two 3,5-dihydroxybenzoate molecules within each four-molecule hydrogen bonded ring (Figure 6.29). The water molecule acts as a hydrogen bond donor in both O-H \cdots O hydrogen bonds to the carboxylate group of the 3,5-dihydroxybenzoic acid molecules. These hydrogen bonds are of moderate strength and are of O \cdots O distances 2.6912(15) Å and 2.7267(15) Å. This water molecule is then involved in two other O-H \cdots O hydrogen bonds, this time acting as a hydrogen bond acceptor from hydroxyl groups of two 3,5-dihydroxybenzoate molecules in a parallel plane (Figure 6.30). These hydrogen bonds are of moderate strength and have O \cdots O distances of 2.7063(15) Å and 2.7387(15) Å, respectively.

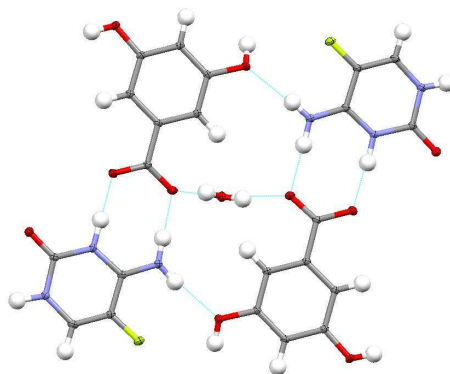


Figure 6.29. Four molecule hydrogen bonded ring with bridging water molecule in the 5-fluorocytosinium 3,5-dihydroxybenzoate hydrate molecular complex.

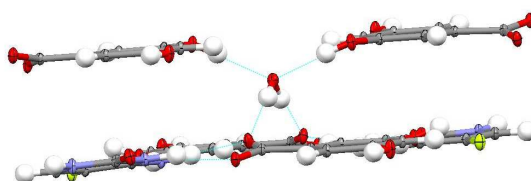


Figure 6.30. Water molecule linking layers in the 5-fluorocytosinium 3,5-dihydroxybenzoate hydrate molecular complex.

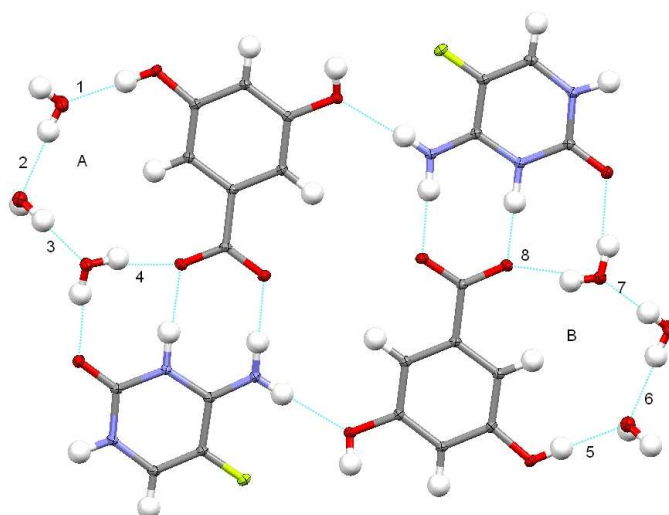


Figure 6.31. Rings formed with water molecules and 3,5-dihydroxybenzoate molecules in the 5-fluorocytosinium 3,5-dihydroxybenzoate hydrate molecular complex.

Within the layers, there are two similar four-molecule rings formed involving one 3,5-dihydroxybenzoate molecule and three waters (Figure 6.31). The orientation of the water molecules in each ring is different. The first ring starts with an O-H...O hydrogen bond from a hydroxy group of a 3,5-dihydroxybenzoate molecule and an oxygen atom of a nearby water molecule. This hydrogen bond is of moderate strength and has an O...O distance of 2.5986(17) Å. The next hydrogen bond in the ring involves two water molecules and is another O-H...O bond with an O...O distance of 2.747(2) Å (2). The next bond in the ring once again involves two water molecules and is another O-H...O bond with an O...O distance of 2.668(2) Å (3). The final bond to complete the ring involves a water molecule and the oxygen atom of a carboxylate group of the same 3,5-dihydroxybenzoate molecule. This bond is also an O-H...O hydrogen bond and is of moderate strength with an O...O distance of 2.731(2) Å (4). This final water molecule also forms a hydrogen bond to a 5-fluorocytosinium molecule creating a smaller three molecule hydrogen bonded ring. This hydrogen bond is of moderate strength and has an O...O distance of 2.823(2) Å.

The second ring is made up in a very similar manner to the ring described previously. However, it is apparent from Figure 6.31, that the orientation of two of the water molecules in each ring differs. The second ring consists of four hydrogen bonds just like the first ring. The first hydrogen bond comes from the hydroxyl group of the benzoic acid and the oxygen atom of the nearby water molecule. This hydrogen bond is of moderate strength and has an O...O distance of 2.690 (2) Å. The second hydrogen bond involves the oxygen atom of the water involved in the first hydrogen bond and the hydrogen atom of the nearby water molecule. This bond is of moderate strength and has an O...O distance of 2.747 (2) Å. The next hydrogen bond involves the other hydrogen bond of the last water molecule

and the oxygen atom of a nearby water molecule. This bond is of moderate strength and has an O...O distance of 2.717 (2) Å. This final water molecule is involved in two more hydrogen bonds with one involving the oxygen atom of the nearby 5-fluorocytosinium molecule and the other involves the oxygen atom of the nearby carboxylate group of the 3,5-dihydroxybenzoate molecule. These bonds are of moderate strength and have O...O distances of 2.848 (2) Å and 2.734 (2) Å, respectively.

Figure 6.32 shows that the water molecules help to link the separate layers together. Two of the water molecules sit in the same plane as the 5-fluorocytosinium and 3,5-dihydroxybenzoate plane, whilst the other two water molecules sit in-between the layers. In the global structure there are alternating ways to link the layers, one bridging the 3,5-dihydroxybenzoate molecules in the four membered ring and one bridging through water molecules.

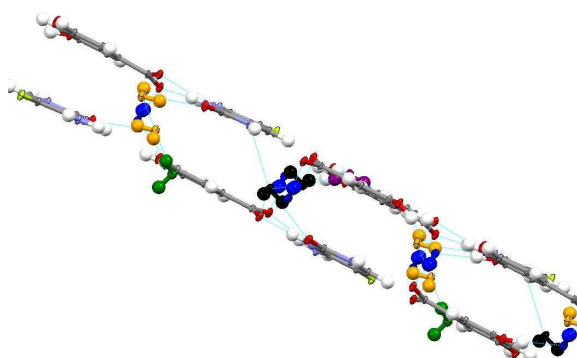


Figure 6.32. The changes in the orientation of the water molecules helping to link the in the 5-fluorocytosinium 3,5-dihydroxybenzoate hydrate molecular complex. Green and purple water molecules in same plane as the chains and the black, blue and orange molecules sat between planes connecting separate chains.

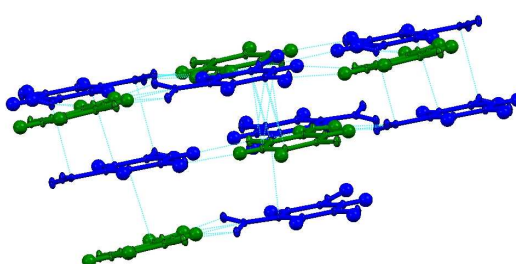


Figure 6.33. Stacking interactions in the 5-fluorocytosinium 3,5-dihydroxybenzoate molecular complex.

This structure contains several stacking interactions but importantly no base stacking interactions (Figure 6.33). Here there are portions of the structure that have a 3,5-dihydroxybenzoate molecule stacked on top of another 3,5-dihydroxybenzoate molecule, and regions where there is a 5-fluorocytosinium stacked on top of 3,5-dihydroxybenzoate molecules. This gives three stacking distances of approximately 3.247 Å, 3.231 Å and 3.161 Å, respectively.

6.5 2:2:1 Molecular Complex of Cytosinium 3,5-Dihydroxybenzoate Hemihydrate

Cytosine and 3,5-dihydroxybenzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.3mm x 0.3mm x 0.3mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 6.5.

Like the fluorinated version of this structure there is 100% proton transfer from the carboxylic acid group of the 3,5-dihydroxybenzoic acid to a cytosine ring nitrogen. Unusually, there is no homo base-pairing in this complex, however, a hydrogen bonded heterodimer between the two components is formed (Figure 6.34); there are two such independent units within the structure. Both these units are constructed of the same hydrogen bonding motif of two N-H...O hydrogen bonds each involving one of the oxygen atoms of the carboxylate group. The cytosine molecule acts as the hydrogen bond donor in both with one hydrogen bond from the amine group and one from a protonated ring nitrogen. In the first, the hydrogen bond involving the amine group has an N...O distance of 2.763(2) Å and the hydrogen bond involving the ring nitrogen has an N...O distance of 2.716(3) Å (Figure 6.34, left, hereon referred to as the symmetric dimer, see below). In the second unit, these hydrogen bonds have N...O distances of 2.794(6) Å and, 2.762(2) Å, respectively (Figure 6.34, right, hereon referred to as the asymmetric dimer, see below).

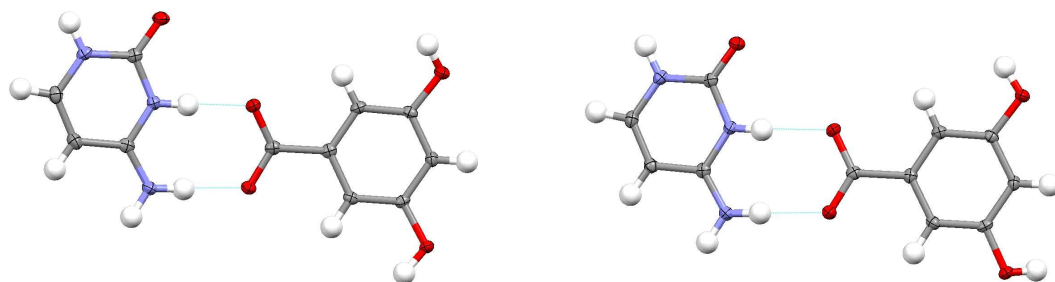


Figure 6.34. Two independent hydrogen bonded rings formed in the molecular complex of cytosinium 3,5-dihydroxybenzoate hemihydrate. Left, the 3,5-dihydroxybenzoate adopts a symmetric configuration, and right, the 3,5-dihydroxybenzoate adopts an asymmetric configuration.

There are two distinctly different configurations for the 3,5-dihydroxybenzoate molecules in the molecular complex, differentiated by the positioning of one of the hydrogen atoms of one of the hydroxyl groups. In one of the 3,5-dihydroxybenzoate molecules, both of the hydrogen atoms are orientated such that without taking bond lengths into account, a line of symmetry can be drawn down the middle of the molecule through the centre of the carboxylate group and perpendicular to the benzene ring (Figure 6.35, left). The second 3,5-dihydroxybenzoate molecule does not have this symmetry due to a different orientation

of the two hydroxyl groups (Figure 6.35, right). This has a dramatic effect on the hydrogen bonding environment around both molecules.

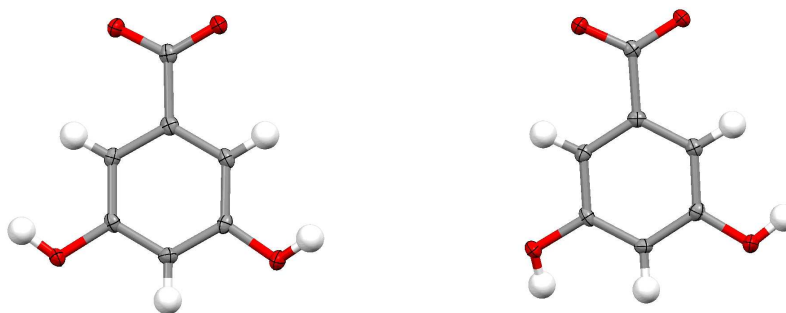


Figure 6.35. Two different 3,5-dihydroxybenzoate molecules in the cytosinium 3,5-dihydroxybenzoate molecular complex.

Two parallel symmetric heterodimers are tethered to one asymmetric dimer cytosine molecule which lies approximately perpendicular to them through N-H...O hydrogen bonds; the first is from a ring nitrogen of the cytosinium in the asymmetric dimer to the carboxylate group of a neighbouring symmetric dimer with an N...O distance of 2.730(2) Å; the second is from a ring nitrogen on a symmetric dimer to the carbonyl group of a cytosinium in the asymmetric dimer with an N...O distance of 2.751(3) Å.

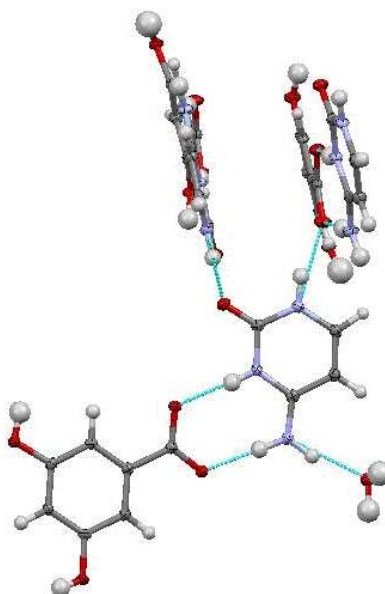


Figure 6.36. Linking of two heterodimers through an N-H...O hydrogen bond from a cytosinium to a 3,5-dihydroxybenzoate molecule.

The cytosinium molecule in the asymmetric dimer also forms a hydrogen bond to a water molecule. This N-H...O hydrogen bond is from the amine group of the cytosinium to the oxygen of a water molecule (Figure 6.36) with an N...O distance of 2.787(4) Å.

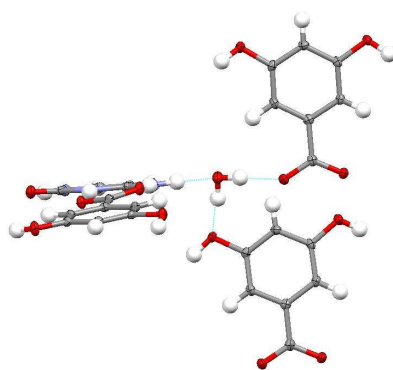


Figure 6.37. Water molecule linking two symmetric 3,5-dihydroxybenzoate molecules to an asymmetric dimer in the cytosinium 3,5-dihydroxybenzoate hemihydrate molecular complex.

This water molecule acts as a hydrogen bond donor in two further hydrogen bonds to symmetric 3,5-dihydroxybenzoate molecules (Figure 6.37). The hydrogen bond for the first of these lies in the same plane as the asymmetric dimer and the other hydrogen bond lies approximately perpendicular to this. The hydrogen bond in the plane is between the water and the carboxylate group of a 3,5-dihydroxybenzoate molecule. This bond is an O \cdots H hydrogen bond and has an O \cdots O distance of 2.750(2) Å and is of moderate strength. This 3,5-dihydroxybenzoate molecule lies approximately perpendicular to the asymmetric dimer. The second hydrogen bond is between the water molecules and one of the oxygen atoms of a hydroxy group of a separate 3,5-dihydroxybenzoate molecule. This 3,5-dihydroxybenzoate molecule is also approximately perpendicular to the asymmetric dimer. This bond is of moderate strength and has an O \cdots O distance of 2.967(2) Å.

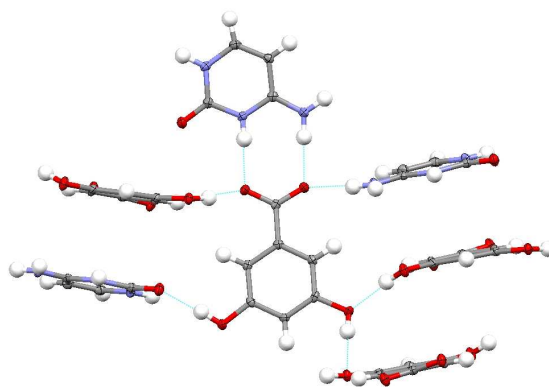


Figure 6.38. Hydrogen bonds involving the asymmetric dimer in the cytosinium 3,5-dihydroxybenzoate hemihydrate molecular complex.

The asymmetric 3,5-dihydroxybenzoate molecule is involved in five hydrogen bonds in addition to the hydrogen-bonded ring with the cytosinium molecule. The two oxygen atoms of the carboxylate group are both involved in hydrogen bonds, one accepting an N-H \cdots O hydrogen bond from the amine group of a cytosinium and the other accepting an O-H \cdots O hydrogen bond from the hydroxyl group of a nearby 3,5-dihydroxybenzoate (Figure 6.38). These bonds are both of moderate strength and have N \cdots O and O \cdots O distances of 2.852(2) Å and 2.617(3) Å, respectively. The hydroxyl group that points

towards the carboxylate group forms an O-H...O hydrogen bond to the oxygen atom of a cytosinium molecule of O...O distance 2.700(2) Å. All these bonds are to molecules which lie approximately perpendicular to the asymmetric dimer (Figure 6.38). The last two hydrogen bonds involving this molecule both involve the hydroxyl group which points away from the carboxylate group of the 3,5-dihydroxybenzoate molecule which acts as a hydrogen bond donor and a hydrogen bond acceptor. The first of these hydrogen bonds is from the hydroxyl group to create an O-H...O hydrogen bond to one of the oxygen atoms of a nearby hydroxyl group of another symmetric 3,5-dihydroxybenzoate molecule. This bond is of moderate strength and has an O...O distance of 2.794(2) Å. The second hydrogen bond is accepted from the hydroxy group of a different symmetric hydroxy group. This hydrogen bond is an O-H...O bond, is of moderate strength and has an O...O distance of 2.872(2) Å. The relative orientations of the two dimers is shown in Figure 6.39.

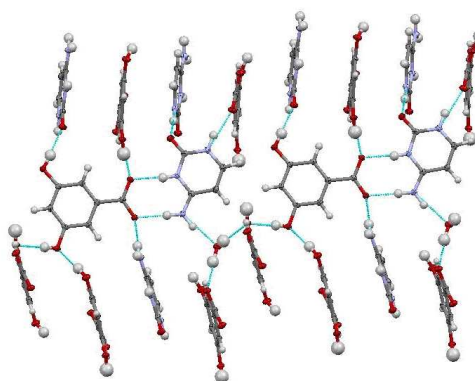


Figure 6.39. The different orientations of the two heterodimers in the cytosinium 3,5-dihydroxybenzoate hemihydrate molecular complex.

The hydrogen bonding in this molecular complex does not result, in this case in a layered structure. There are still several stacking interactions. Figure 6.40 shows that the symmetric 3,5-dihydroxybenzoate molecule (red) stacks on the cytosinium molecule of another symmetric dimer (blue). The stacking distances for these interactions are approximately 2.933 Å and 3.230 Å. There are two other stacking interactions between a cytosinium molecule in an asymmetric dimer (orange) and another asymmetric 3,5-dihydroxybenzoate molecule (green). The stacking distances are approximately 3.247 Å and 3.263 Å.

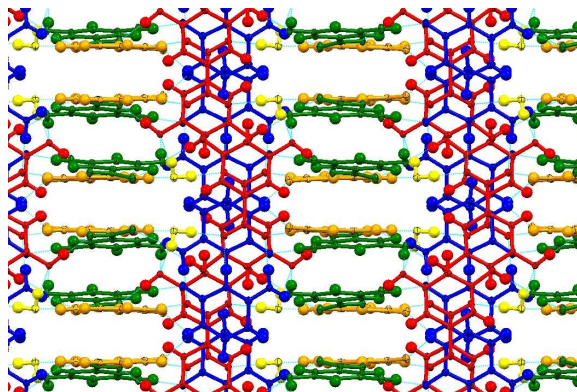


Figure 6.40. Stacking interactions in the cytosinium 3,5-dihydroxybenzoate hemihydrate molecular complex. Red and blue are the 3,5-dihydroxybenzoate and cytosinium molecules in a symmetric dimer, respectively and green and orange are the 3,5-dihydroxybenzoate and cytosinium molecules in an asymmetric dimer, respectively. Water is yellow.

6.6. 1:4:5 Molecular Complex of 5-Fluorouracil and 3,5-Dihydroxybenzoic Acid Hydrate

5-fluorouracil and 3,5-dihydroxybenzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.3mm x 0.3mm x 0.3mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 6.5.

The 5-fluorouracil molecule cannot accept any further hydrogen atoms and so the molecular complex consists entirely of neutral species. The molecular complex of 5-fluorouracil 3,5-dihydroxybenzoic acid hydrate forms in an unusual 1:4:5 ratio. Surprisingly, there is no conventional primary hydrogen bonding motif involving the 5-fluorouracil molecule in this molecular complex. All of the hydrogen bonds involving the 5-fluorouracil involve only water molecules or base-pairing with another 5-fluorouracil through moderate C-H...F hydrogen bonds. There are no direct interactions with 3,5-dihydroxybenzoic acid molecules.

Table 6.1. The hydrogen distances for the hydrogen bonds between 5-fluorouracil molecules and water molecules in the 5-fluorouracil 3,5-dihydroxybenzoic acid hydrate molecular complex. The lettering corresponds to the hydrogen bonds shown in Figure 6.41.

Hydrogen Bond	D...A Bond length
A	2.823(2) Å
B	2.793(3) Å
C	2.725(3) Å
D	2.783(2) Å
E	2.707(3) Å
F	2.693(2) Å

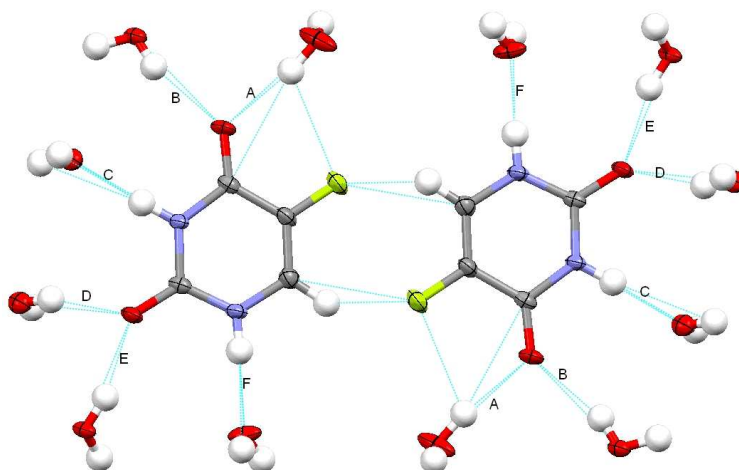


Figure 6.41. Hydrogen bonds involving the 5-fluorouracil molecule in the molecular complex of 5-fluorouracil 3,5-dihydroxybenzoic acid hydrate.

The 5-fluorouracil molecule forms homodimers, this time held together by the relatively weaker C-H...F hydrogen bonds (Figure 6.41). There is an inversion centre in the centre of the resulting hydrogen bonded ring and the hydrogen bonding between 5-fluorouracil molecules resembles a base pairing motif. These two moderate strength hydrogen bonds are symmetrically equivalent and have a C...F distance of 3.138(4) Å. There are a further six hydrogen bonds with water molecules for each 5-fluorouracil molecule, two where the fluorouracil molecule acts as a hydrogen bond donor and four where it acts as a hydrogen bond acceptor (Figure 6.41). All of these bonds are of moderate strength. These water molecules fulfil two roles; the first disrupts the potential pseudo-Hoogsteen base pairing acting as a spacer between the two 5-fluorouracil molecules (Figure 6.42); the second is as

a link between these 5-fluorouracil dimers and the 3,5-dihydroxybenzoic acid molecules. The hydrogen bonds acting as the spacer molecules involve two symmetry equivalent water molecules with two independent hydrogen bonds (C and D, Figure 6.43) in the four molecule hydrogen bonded unit (Figure 6.42).

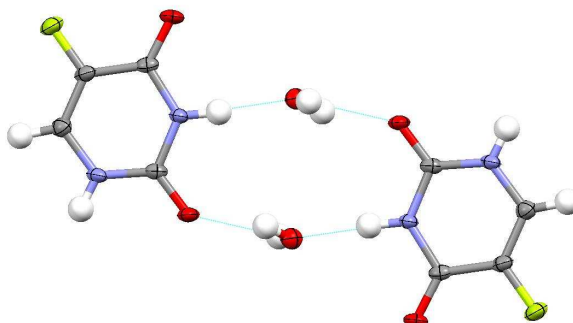


Figure 6.42. Water molecules disrupting the type B base-pairing in the molecular complex of 5-fluorouracil 3,5-dihydroxybenzoic acid hydrate.

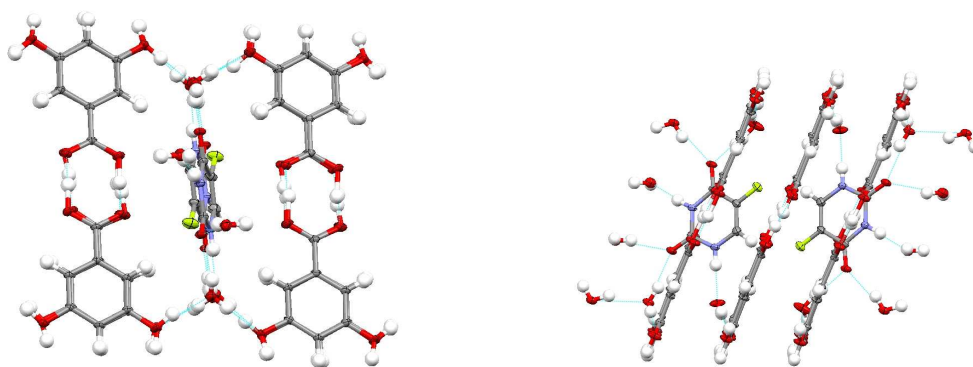


Figure 6.43. 5-fluorouracil surrounded by a ring of 3,5-dihydroxybenzoic acid and water molecules in molecular complex of 5-fluorouracil and 3,5-dihydroxybenzoic acid hydrate.

Pairs of 3,5-dihydroxybenzoic acid molecules form homodimers characteristic of benzoic acid molecules, held together by two O-H...O hydrogen bonds of moderate strength. There are three independent dimers formed, two of which form between the same molecule related by inversion (hydrogen bonds H and K in Figure 6.44), and one which forms between two independent 3,5-dihydroxybenzoic acid molecules (hydrogen bonds N and O in Figure 6.44). The hydrogen bonds lengths are summarised in Table 6.2.

The 5-fluorouracil molecules are surrounded by two independent 3,5-dihydroxybenzoic acid and water rings (Figure 6.43, Figure 6.44). Each ring is comprised of four 3,5-dihydroxybenzoic acid molecules and two water molecules. Each is held together by eight hydrogen bonds, four of which are part of the 3,5-dihydroxybenzoic acid dimers; one of the hydrogen bonded rings has an inversion centre located in the centre of the ring (Figure 6.44, left).

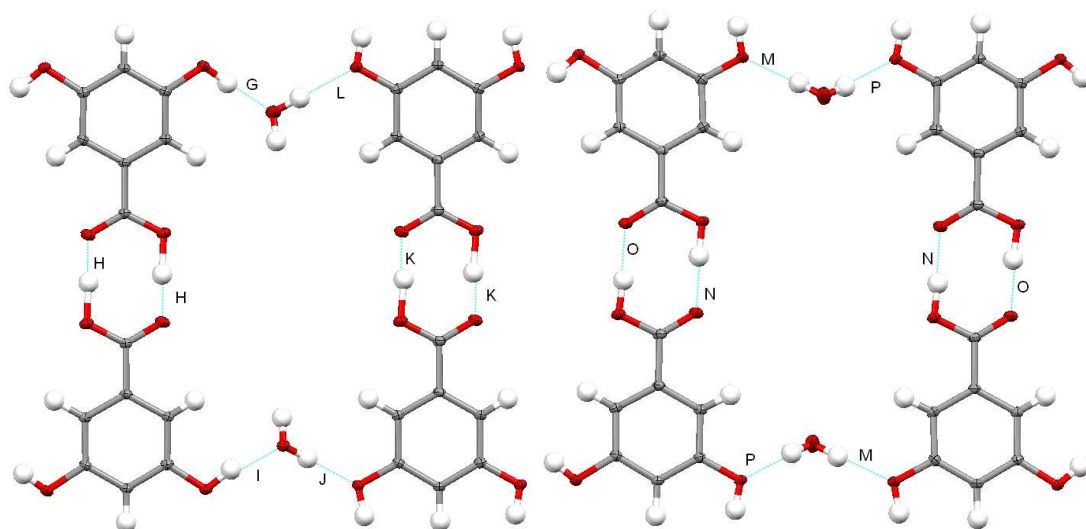


Figure 6.44 Ring 1 (left) and ring 2 (right) made up of 3,5-dihydroxybenzoic acid and water molecules in the molecular complex of 5-fluorouracil and 3,5-dihydroxybenzoic acid hydrate.

Table 6.2. Hydrogen bond lengths of the 3,5-dihydroxybenzoic acid dimers and the hydrogen bonded rings in the molecular complex of 5-fluorouracil 3,5-dihydroxybenzoic acid hydrate. Bond labeling corresponds to those indicated in Figure 6.44.

Hydrogen bonds in ring 1	D...A bond length	Hydrogen bonds in ring 2	D...A bond length
G	2.603(2) Å	M	2.762(3) Å
H	2.613(3) Å	N	2.655(2) Å
I	2.644(3) Å	O	2.650(3) Å
J	2.767(2) Å	P	2.816(2) Å
K	2.631(3) Å		
L	2.777(2) Å		

The base paired 5-fluorouracil is surrounded by four rings made up of benzoic acid molecules and water molecules. The order alternates between ring 1 and ring 2. The rings are connected to the base pair via moderate strength hydrogen bonds (Figure 6.45). The connections from ring 1 to the base pair are from the hydrogen atom of the water molecules in the ring and involves the two oxygen atoms of the 5-fluorouracil. These hydrogen bonds have an O...O distance of 2.707(2) Å (Q) and 2.823(2) Å (R). The connection between ring 2 and the base pair is via two equivalent hydrogen bonds between the one ring and two separate 5-fluorouracil molecules. The connection is between the oxygen atom of a water molecule in ring 2 and a protonated heteroatom in the 5-

fluorouracil ring (Figure 6.5, right). This occurs twice and both bonds are of moderate strength and both have an N...O distance of 2.693(2) Å.

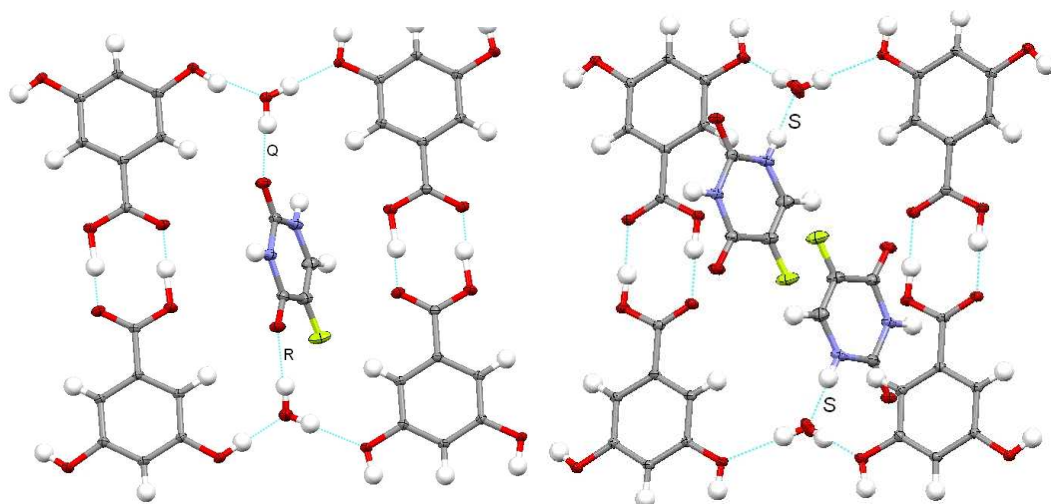


Figure 6.45. Ring 1 with link to the base pair (left) and ring 2 with link to the base pair (right) via water molecules in molecular complex of 5-fluorouracil and 3,5-dihydroxybenzoic acid hydrate.

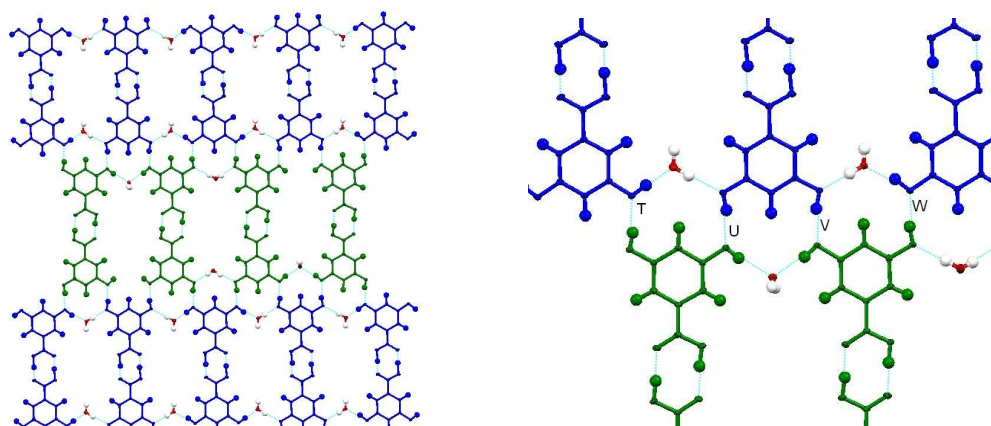


Figure 6.46. Hydrogen bonds linking separate rings in the same plane in the molecular complex of 5-fluorouracil and 3,5-dihydroxybenzoic acid hydrate.

The rings are linked four unique hydrogen bonds. All of the hydrogen bonds here are of moderate strength and involve hydrogen bonds between hydroxyl groups of the 3,5-dihydroxybenzoic acid molecules. There are no other connections between molecules within the planes. These hydrogen bonds have O...O distances of 2.736(2) Å (T), 2.740(2) Å (U), 2.742(2) Å (V) and 2.698(2) Å (W) (Figure 6.46).

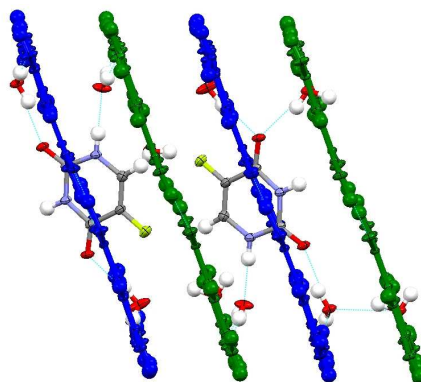


Figure 6.47. Stacking interactions in the molecular complex of 5-fluorouracil and 3,5-dihydroxybenzoic acid hydrate.

There are also stacking interactions between the 3,5-dihydroxybenzoic acid framework layers. The approximate spacings between these layers are 3.388 Å and 3.335 Å.

6.7. 1:1 Molecular Complex of 5-Fluorocytosine and 3,5-Dinitrobenzoic Acid

5-Fluorocytosine and 3,5-dinitrobenzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.3mm x 0.3mm x 0.3mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 6.5.

This structure behaves in a different manner to the other structures in the series including some unusual hydrogen transfer effects. In the other structures that contain cytosine reported in this thesis, it is common for there to be hydrogen transfer from the carboxylic acid group of the benzoic acid to the unprotonated nitrogen of the cytosine. This molecular complex forms in a 1:1 ratio in space group $C2/c$, however only partial hydrogen transfer occurs from the carboxylic acid group of the 3,5-dinitrobenzoic acid to the unprotonated heteroatom in the 5-fluorocytosine ring. Instead of there being four different molecular species within the asymmetric unit (one 5-fluorocytosine, one 5-fluorocytosinium, one 3,5-dinitrobenzoic acid and one 3,5-dinitrobenzoate), there are only two, one partially protonated 5-fluorocytosine molecule and one partially deprotonated 3,5-dinitrobenzoic acid molecule.

Two homodimers are formed. A consequence of the partial hydrogen transfer is the formation of a pseudo-Watson-Crick base pairing motif between two 5-fluorocytosine molecules (Figure 6.48, left). An inversion centre is located in the middle of the central hydrogen bond which results in a single hydrogen atom being effectively shared between

the two molecules in this unit. There are two possible models for this hydrogen atom: it could either lie on the inversion centre and therefore adopt a central position within the hydrogen bond, or be disordered over two positions within the hydrogen bond. Both models give a good fit to the data. The D...A distances are given in Table 6.3.

The second homodimer is between two 3,5-dinitrobenzoic acid molecules, however these are only held together by a single hydrogen atom rather than the more conventional double hydrogen bond in standard benzoic acid dimer (Figure 6.48, right). This is a consequence of the partial deprotonation of the 3,5-dinitrobenzoic acid molecule. The O...O distance for this hydrogen bond is 2.466(2) Å, making it a strong hydrogen bond. There are again two possible models for the hydrogen atom shared between the two molecules. Firstly it could be located on the inversion centre in the middle of the O-H...O hydrogen bond, thus being shared equally between the two molecules. Secondly, it could adopt a 50:50 disordered position across the hydrogen bond. Both models again provide a good fit to the data. The carboxylic acid group has bond lengths of 1.234(2) Å and 1.282(2) Å and when these are compared to the distances observed in the pure structure of 3,5-dinitrobenzoic acid [136] of 1.250 Å and 1.274 Å it is clear to see that one length has shortened and the other length has lengthened. However, the changes are not dramatic and these may be difficult to compare due to the disorder present in the structure. The protonated oxygen atom of the carboxylic acid group is still showing single bond character, whilst the non-protonated oxygen atom is close to the expected length for a double bond. Neutron diffraction would thus be required to characterise fully the hydrogen behaviour within both of these hydrogen bonds. Unfortunately, attempts to grow neutron sized crystals have to-date, been unsuccessful.

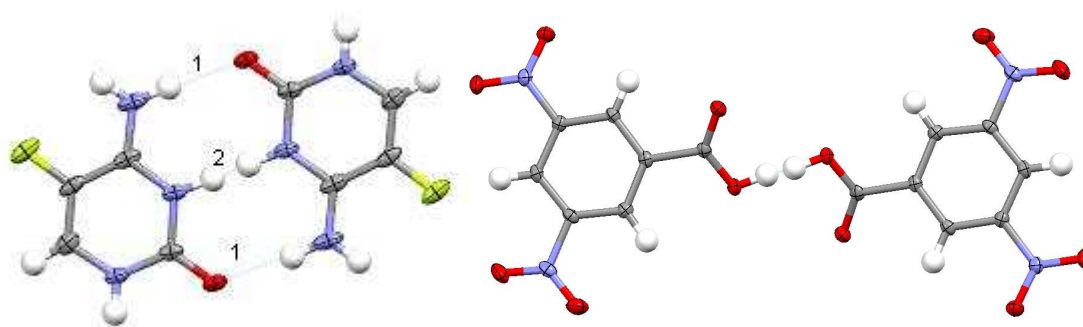


Figure 6.48 Pseudo-Watson-Crick unit (left) and the 3,5-dinitrobenzoic acid dimer (right) in the molecular complex of 5-fluorocytosine and 3,5-dinitrobenzoic acid.

Table 6.3. The hydrogen bond distances in the pseudo-Watson-Crick unit in the 5-fluorocytosine 3,5-dinitrobenzoic acid molecular complex.

Hydrogen Bond	D...A Bond length
1	2.814(2) Å
2	2.790(2) Å
1	2.814(2) Å

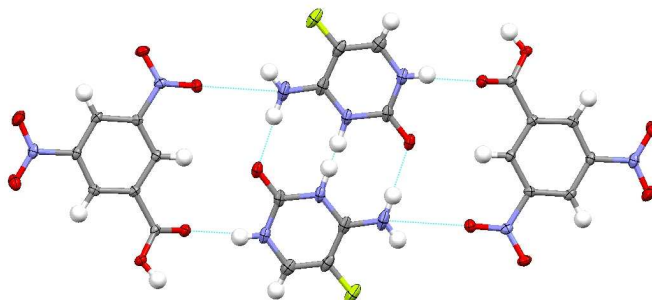


Figure 6.49. Four molecule building block in the molecular complex of 5-fluorocytosine and 3,5-dinitrobenzoic acid

The structure contains a basic building block that is made up of four molecules and involves seven hydrogen bonds. The central unit comprising three of these hydrogen bonds is the pseudo-Watson-Crick base pair. On either side of this there is a 3,5-dinitrobenzoic acid molecule tied to the unit via two hydrogen bonds. The first of these hydrogen bonds is from the amine group of the 5-fluorocytosine molecule to an oxygen atom of one of the nitro groups of the 3,5-dinitrobenzoic acid molecule. This hydrogen bond is an N-H...O hydrogen bond and is of moderate strength with an N...O distance of 2.951(2) Å. The second hydrogen bond that completes the link between the 3,5-dinitrobenzoic acid and the base pair involves the normally protonated ring nitrogen of the cytosine and the non-protonated oxygen atom of the carboxyl group of the benzoic acid molecule. This bond is of moderate strength and has an N...O distance of 2.710(2) Å. This unit is symmetrical and so these hydrogen bonds are formed on both sides of the base pair. These four molecule units assemble into chains, held together by the benzoic acid dimer (black box in Figure 6.50).

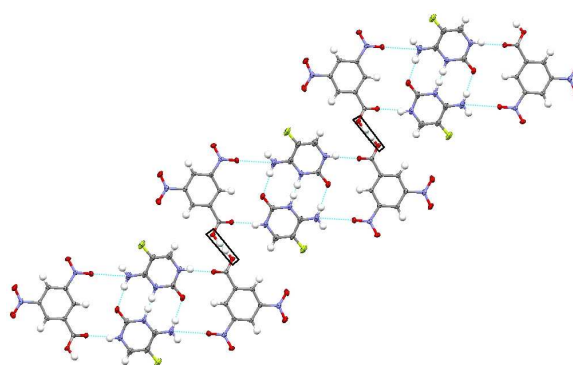


Figure 6.50. 3,5-dinitrobenzoic acid dimer linking separate four molecule building blocks together in the molecular complex of 5-fluorocytosine and 3,5-dinitrobenzoic acid.

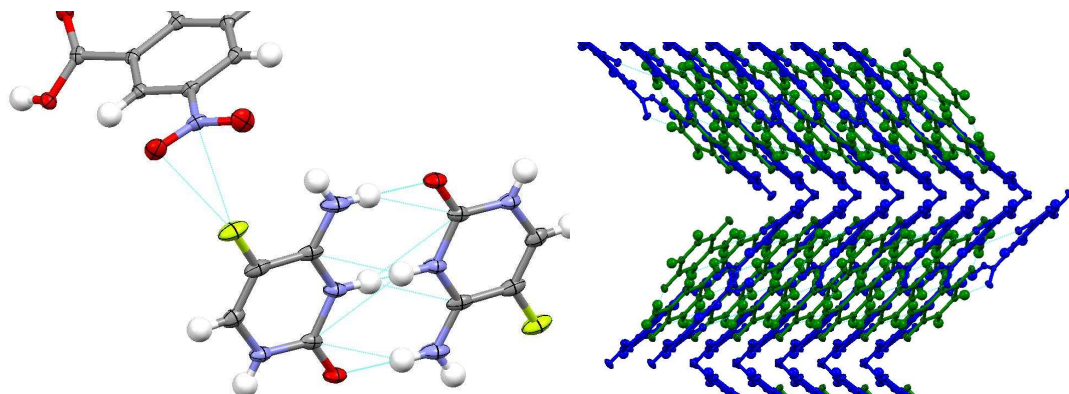


Figure 6.51. Fluorine interactions (left) and the herringbone structure produced (right) in the molecular complex of 5-fluorocytosine and 3,5-dinitrobenzoic acid.

There is a short F...N contact to the nitrogen atom of one of the nitro groups of the 3,5-dinitrobenzoic acid molecule (Figure 6.51, left). This interaction has an F...N distance of 2.880(3) Å and this is significant when compared to the sum of the van der Waals radii of the two atoms (3.02 Å). This fluorine interaction helps to link chains that run perpendicular to each other. This forms a herring-bone structure (Figure 6.51, right).

There are stacking interactions between cytosine molecules stacked on top of each other and 3,5-dinitrobenzoic acid molecules stacked on top of each other. The base stacking distances are approximately 3.187 Å and 3.317 Å, respectively. These stacking interactions are the only direct links between the molecules of layers stacked one on top of the other.

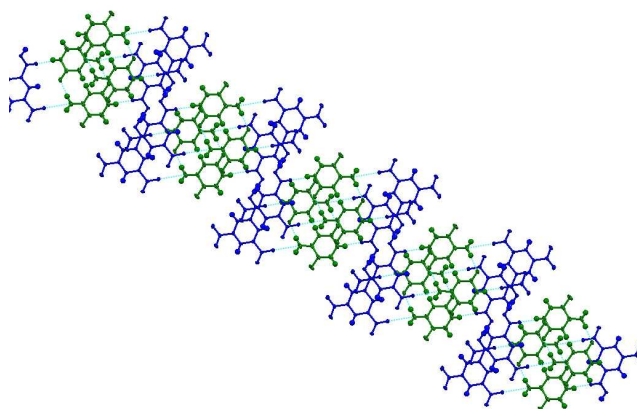


Figure 6.52 Stacking interactions in the molecular complex of 5-fluorocytosine and 3,5-dinitrobenzoic acid.

6.8. Structural Trends in Molecular Complexes with of Cytosine and 5-Fluorocytosine Disubstituted Hydroxybenzoic Acids

pK_a values can be a useful indicator of the likely proton transfer effects in molecular complexes (Table 6.4) Section 1.6. All of the molecular complexes show total hydrogen transfer in this series except for the 2,6-dihydroxybenzoic acid and cytosine hydrate molecular complex.

Table 6.4. Summary of the molecular complexes of cytosine and 5-fluorocytosine reported in this chapter.

Molecular complex	ΔpK_a	Hydrogen transfer with respect to cytosine molecules	Crystallisation ratio
Cytosinium 2,4-dihydroxybenzoate hydrate Form I	1.49	100%	1:1:1
Cytosinium 2,4-dihydroxybenzoate hydrate Form II	1.49	100%	1:1:1
5-Fluorocytosinium 2,6-dihydroxybenzoate Form I	1.96	100%	1:1
5-Fluorocytosinium 2,6-dihydroxybenzoate Form II	1.96	100%	1:1
Cytosine 2,6-dihydroxybenzoic acid hydrate	3.30	50%	1:2:1
5-fluorocytosine 3,5-dihydroxybenzoate hydrate	-0.78	100%	2:2:5
Cytosinium 3,5-dihydroxybenzoate hydrate	0.56	100%	2:2:1

6.8.1 Molecular Complexes with 2,4-Dihydroxybenzoic Acid

There are two structures that involve 2,4-dihydroxybenzoic acid; these are polymorphs with cytosine and both show total hydrogen transfer. The ΔpK_a value for both is therefore 1.49 (Table 6.4), which falls into the region where there is a possibility of both salt and co-crystal formation. These structures contain total hydrogen transfer so, except for the water molecules, there are no neutral molecules present in the structure.

6.8.2 Molecular Complexes of 2,6-Dihydroxybenzoic Acid

The molecular complexes reported in this chapter with 2,6-dihydroxybenzoic acid are with both cytosine and 5-fluorocytosine. The difference in pK_a is quite large in this case as the pK_a of 2,6-dihydroxybenzoic acid is 1.3. These structures follow the “rule of three” that when the ΔpK_a value is greater than 3, as is the case for the cytosine molecular complex (Table 6.4), proton transfer is the most likely outcome. This structure contains a water molecule and mixed protonation states of cytosine molecules. The rule has been successfully followed even though there is a water molecule present in the structure. This structure falls into the region where a salt is expected to form and the structure follows the rules laid out.

The remaining structures, the molecular complexes of 5-fluorocytosine with 2,6-dihydroxybenzoic acid polymorphs, these structures have ΔpK_a values of 1.96, in the continuum region for proton transfer prediction. This structure contains total hydrogen transfer. The crystallisation ratio is 1:1 and the rules have been successfully followed here.

6.8.3 Molecular Complexes of 3,5-Dihydroxybenzoic Acid

In the molecular complexes with 3,5-dihydroxybenzoic acid, the hydrogen transfer is total. The ΔpK_a values fall into the region where it is difficult to predict whether proton transfer will occur. The pK_a of 3,5-dihydroxybenzoic acid is 4.04 and this means that in the example of 5-fluorocytosine and 3,5-dihydroxybenzoic acid hydrogen transfer is unfavourable and pK_a matching would indicate the formation of a neutral molecular complex; however, the structure shows total transfer. The proton transfer in this case may be influenced by the presence of water. In this pair of structures the ratio of crystallisation is 2:2:5, but the quantity of water is varied with the cytosine molecular complex containing five times less water than the cytosine molecular complex. Despite this, the same basic building blocks are observed in both, the hydrogen bonded heterodimer.

6.8.4. Hydrogen Bonding Motifs

The hydrogen bonding motifs formed in this series of structures are consistent with those found in the other chapters in this thesis. The hydrogen transfer is the primary factor in determining which hydrogen bonding motif is adopted by the structure. All but one of the structures contains the hydrogen bonded heterodimer. This is the cytosine and 2,6-dihydroxybenzoic acid molecular complex and this only shows partial hydrogen transfer (there are both neutral and charged cytosine molecules present in the structure). This means the pseudo-Watson-Crick bonding motif is formed in this complex. The pseudo-Watson-Crick base-pairing is only possible in structure containing cytosine when there are both charged and neutral molecules in the structure.

The two polymorphs of 5-fluorocytosinium 2,6-dihydroxybenzoate hydrate and the molecular complex of cytosinium 3,5-dihydroxybenzoate hydrate form the hydrogen bonded heterodimer but the connecting bonds to nearby pairs is not through another base pairing motif, but through single hydrogen bonds. The other three structures containing the hydrogen bonded ring motif are linked to one another via the type B pseudo-Hoogsteen bonding motif. These are the two polymorphs of cytosinium 2,4-dihydroxybenzoate hydrate and the molecular complex of 5-fluorocytosinium 3,5-dihydroxybenzoate hydrate. This helps to form four molecule building blocks in each of the structures and these building blocks are linked together via single hydrogen bonds or via water molecules.

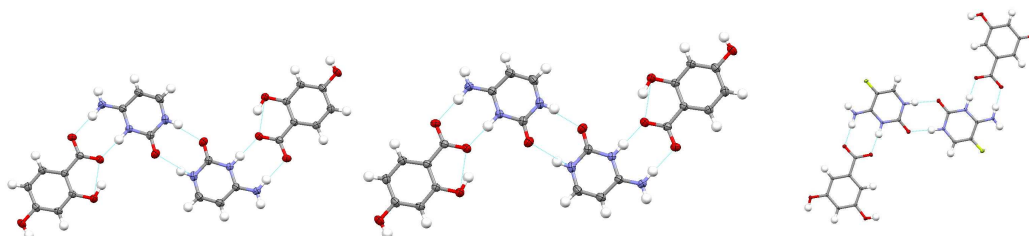


Figure 6.53. Similar building blocks found in the molecular complexes of cytosinium 2,4-dihydroxybenzoate hydrate Form I (left), Form II (middle) and 5-fluorocytosinium 3,5-dihydroxybenzoate hydrate (right).

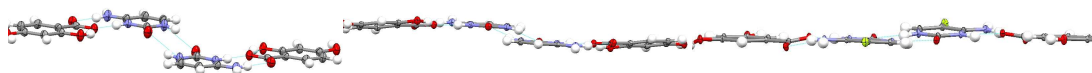


Figure 6.54. Steps in the four molecule building blocks found in the molecular complexes of cytosine and 2,4-dihydroxybenzoic acid hydrate Form I (left), Form II (middle), and 5-fluorocytosine and 3,5-dihydroxybenzoic acid hydrate (right).

In these four molecule building blocks there is the same hydrogen bonding in each of the structures, however, some of these units are not planar and show a step. In the molecular complex of cytosinium 2,4-dihydroxybenzoate Form I, the step is the most pronounced

whilst Form II is much less significant. The molecular complex of 5-fluorocytosinium and 3,5-dihydroxybenzoate does not contain a noticeable step.

Table 6.5. Crystallographic data for the molecular complexes reported in Chapter 6.

Compound	5-fluorocytosinium 2,6-dihydroxybenzoate Form I	5-fluorocytosinium 2,6-dihydroxybenzoate Form II	Cytosine and 2,6-dihydroxybenzoic acid Hydrate	Cytosine and 2,4-dihydroxybenzoic acid Hydrate Form I	Cytosine and 2,4-dihydroxybenzoic acid Hydrate Form II	5-Fluorocytosine and 3,5-dihydroxybenzoic acid Hydrate	Cytosine and 3,5-dihydroxybenzoic acid hydrate	5-fluorouracil and 3,5-dihydroxybenzoic acid hydrate	5-fluorocytosine and 3,5-dinitrobenzoic acid
Formula	C11 H10 F1 N3 O5	C11 H10 F1 N3 O5	C15 H18 N6 O7	C11 N3 O6 H13	C11 N3 O6 H13	C22 O15 N6 F2 H30	C22 O11 N6 H24	C32 O23 F1 N2 H37	C11 H9 N5 O7 F1
Crystallisation Conditions	Methanol, Room Temperature	Methanol, Room Temperature	Methanol, Room Temperature	Methanol, Room Temperature	Methanol, Room Temperature	Methanol, Room Temperature	Methanol, Room Temperature	Methanol, Room Temperature	Methanol, Room Temperature
Molecular weight / gmol⁻¹	283.21	283.21	394.34	283.2	283.2	656.50	548.46	836.63	342.2
Temperature (K)	100 K	100 K	100 K	100 K	100 K	100 K	100 K	100 K	100 K
Space Group	P2 ₁ /c	P2 ₁ /c	P2 ₁ /n	P-1	P2 ₁ /c	P-1	Pbca	P-1	C2/c
<i>a</i> (Å)	13.248(3)	7.1046(4)	8.346(1)	7.0010(4)	12.991(3)	7.1827(3)	25.1889(2)	10.2429(6)	28.650(8)
<i>b</i> (Å)	6.945(2)	13.9021(8)	20.395(2)	8.6817(5)	6.530(1)	13.4398(6)	13.6500(1)	13.1654(7)	4.820(1)
<i>c</i> (Å)	15.986(2)	12.9934(8)	10.101(1)	11.1981(6)	16.709(3)	15.4909(7)	13.4754(1)	14.7630(8)	21.682(6)
α (°)	90	90	90	90.919(3)	90	109.268(3)	90	93.532(3)	90
β (°)	125.37(1)	116.780(4)	95.826(7)	98.377(3)	120.46(1)	93.297(3)	90	108.103(3)	118.353(8)
γ (°)	90	90	90	112.842(3)	90	102.729(3)	90	104.398(3)	90
Volume (Å³)	1197(1)	1145.7(3)	1710.5(1)	618.6(1)	1221.8(1)	1363.10(22)	4633.23(1)	1811.87(29)	2634.81(13)
Z	4	4	4	2	4	2	8	2	8
θ range (°)	3.1-27.5	2.3-30.5	2.0-34.6	1.8-30.1	3.3-27.5	1.4-33.3	1.6-27.5	1.5-34.25	1.6-20.4
Completeness	99.90%	99.70%	98.90%	98.60%	98.40%	99.10%	99.90%	99.30%	99.10%
Reflections Collected	15201	38582	90833	35429	14444	85995	71875	94191	24775
Independent	2740	3490	7206	3604	2766	10395	5303	10099	1295
Refln (obs.I>2 σ (I))	2005	2790	5542	2005	1892	7932	3665	6650	1032
<i>R</i>_{int}	0.0440	0.0591	0.0475	0.0396	0.0542	0.0436	0.0858	0.0530	0.1138
Parameters	221	221	325	233	233	526	448	671	247
Goof on F²	1.096	1.044	1.036	1.057	1.010	1.096	1.022	1.079	1.073
<i>R</i>₁ (Observed)	0.037	0.041	0.043	0.042	0.045	0.054	0.043	0.061	0.036
<i>R</i>₁ (all)	0.056	0.056	0.061	0.055	0.073	0.073	0.082	0.095	0.056
<i>wR</i>₂ (all)	0.097	0.120	0.131	0.137	0.122	0.172	0.098	0.231	0.100

Chapter 7 Additional Molecular Complexes of Cytosine

A number of additional molecular complexes of cytosine have also been obtained during this project. The discussion here will be limited to the hydrogen bonding motifs obtained and how these fit into the wider context of the other structures reported in this thesis.

7.1. Molecular Complexes of Cytosine with the Isomers Isonicotinic Acid and Nicotinic Acid

7.1.1 2:1:2 Cytosine and Isonicotinic acid Hydrate

Cytosine and isonicotinic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless plates and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.4mm by 0.1mm by 0.3mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 7.6, at the end of this Chapter.

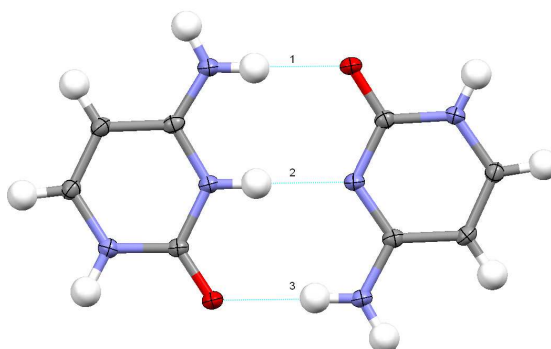


Figure 7.1 Pseudo-Watson-Crick base pair motif in the molecular complex of cytosine isonicotinic acid hydrate

The molecular complex of cytosine isonicotinic acid hydrate crystallises in a 2:1:2 ratio. There is 100% hydrogen transfer from the carboxylic acid group of the isonicotinic acid molecule to one of the ring nitrogens of one cytosine molecule. This results in a deprotonated isonicotinic acid (isonicotinate), a protonated cytosine (cytosinium), a neutral cytosine and two water molecules being present in the asymmetric unit. The 2:1 ratio of cytosine to isonicotinic acid results in the formation of a pseudo-Watson-Crick base pair motif between the protonated and neutral cytosine molecules (Figure 7.1, Table 7.1).

Table 7.1 Hydrogen bonds of the pseudo-Watson-Crick base pair in the molecular complex of cytosine isonicotinic acid hydrate

Hydrogen Bond	D---A Distance
1. O---H-N	2.791(2) Å
2. N---H-N	2.857(2) Å
3. N-H---O	2.915(2) Å

Chains of base pairs are formed where one base pair is connected to another base pair through two additional single N-H...O hydrogen bonds which create a four molecule $R_4^2(8)$ hydrogen bonded ring; two sides of the resultant diamond are also part of the pseudo-Watson-Crick base pair motif (Figure 7.2). These linking hydrogen bonds are of N...O distances of 2.814(2) and 2.834(2)Å, respectively.

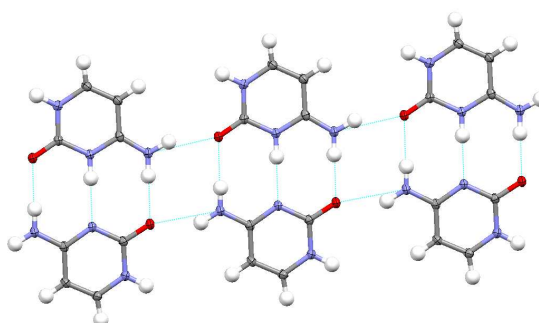


Figure 7.2 Base-paired chain of cytosine molecules with alternating pseudo-Watson-Crick and four molecule hydrogen bonded rings in the molecular complex of cytosine isonicotinic acid hydrate

There are further hydrogen bonds from the cytosine molecules on both sides of the chain; on the cytosinium side, there are two hydrogen bonds, one moderate and one weak to the isonicotinate molecule, and on the neutral side of the chain, there are single hydrogen bonds to water molecules (Figure 7.3). The hydrogen bonds involving the cytosinium molecules are N-H...O hydrogen bonds between a protonated ring nitrogen on the cytosine ring and the carboxylate group of the isonicotinic acid and a C-H...O hydrogen bond from an aromatic carbon to the other oxygen of the carboxylate group. These hydrogen bonds have N...O and C...O distances of 2.679(2) Å and 3.056(2) Å, respectively. The N-H...O hydrogen bond involving the neutral cytosine molecule is between the protonated ring nitrogen and the oxygen of the water molecule. This hydrogen bond is of moderate strength and is of N...O distance of 2.738(2) Å.

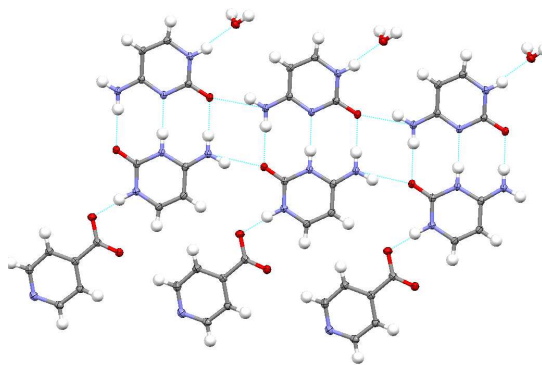


Figure 7.3 Water and isonicotinic acid molecules tied to the cytosine chain via hydrogen bonds in the molecular complex of cytosine isonicotinic acid hydrate

The water molecules are key to both the extension of these cytosine and isonicotinic acid units into layers and in the connection of different layers; this is mediated by two further hydrogen bonds where the water molecule acts as a hydrogen bond donor. The first of these is an O-H...O hydrogen bond to an oxygen atom of the carboxylate group of an isonicotinic acid molecule in a parallel plane. This hydrogen bond has an O...O distance of 2.780(2) Å and is of moderate strength. The other hydrogen bond is to a second independent water molecule within the same layer and these two water molecules together connect two cytosine chains.

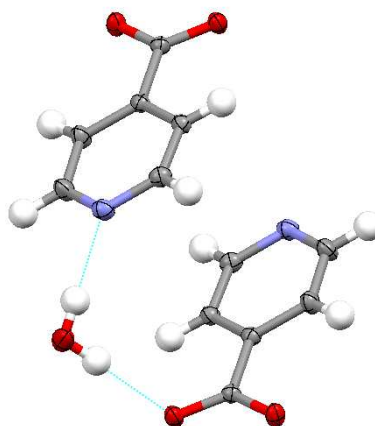


Figure 7.4 Water molecule linking two isonicotinic acid molecules via hydrogen bonding in the molecular complex of cytosine isonicotinic acid hydrate

There are also moderate strength hydrogen bond from a water molecule to two separate isonicotinic acid molecules. The first of these hydrogen bonds is to the ring nitrogen of an isonicotinic acid molecule in the same plane with an O...N distance of 2.857(2) Å and is of moderate strength. The second hydrogen bond is to the oxygen of an isonicotinic acid molecule in a parallel plane by a moderate strength O-H...O hydrogen bond of O...O distance of 2.768(2) Å (Figure 7.4). In this structure, the water and isonicotinate molecules combine to link chains of cytosine molecules together into layers. This is similar to that

seen in other structures reported in this work (3-hydroxybenzoic acid and cytosine hydrate, 5-fluorocytosine and 5-fluorosalicylic acid hydrate and cytosine and 5-fluorosalicylic acid hydrate).

In a similar way to other structures showing these chains of pseudo-Watson-Crick cytosine dimers, (e.g. benzoic acid and cytosine), the chains stack above each other (Figure 7.5, 13ft). Again, the neutral cytosine side of the chain stacks on top of the charged side of another cytosine chain and vice versa. The planes are staggered relative to one another with an approximate base stacking distance between the planes of 3.158 Å (Figure 7.5, right). There are also $\pi\cdots\pi$ stacking interactions between the isonicotinate molecules (Figure 7.6). The isonicotinate molecules, much like the cytosine molecules, do not stack directly on top of each other and adopt a staggered stacking arrangement. The stacking interaction here is of an approximate distance of 3.323 Å.

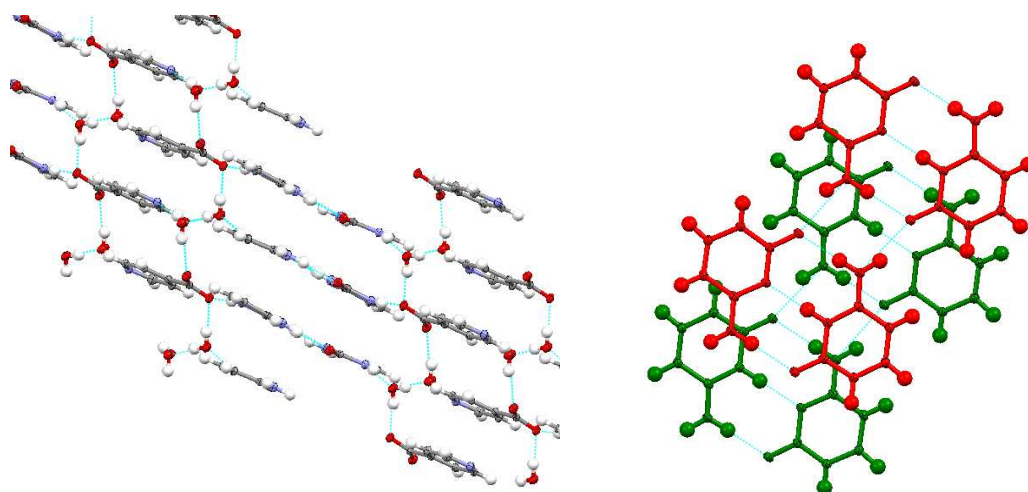


Figure 7.5 Left, layers connected through hydrogen bonding involving the water and right, the staggered stacking of cytosine chains in the molecular complex of cytosine isonicotinic acid hydrate

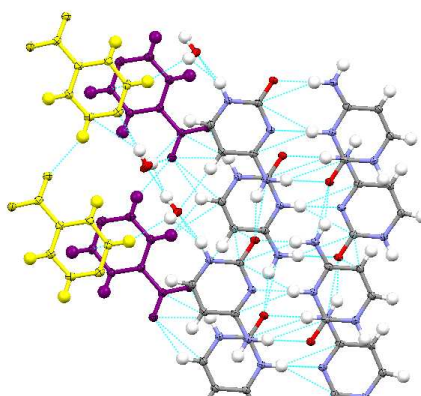


Figure 7.6 Stacking of the isonicotinic acid molecules (yellow and purple) in the molecular complex of cytosine isonicotinic acid hydrate

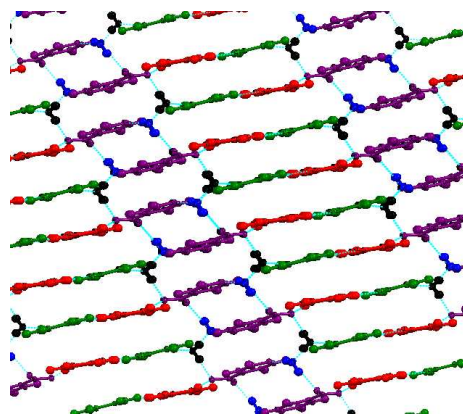


Figure 7.7 Colour coding of the four independent molecules in the molecular complex of cytosine isonicotinic acid hydrate to show the overall packing. One water molecule (blue) another water molecule (black), isonicotinic acid molecule (purple), neutral cytosine (red) and charged cytosine (green)

The overall packing can be seen in Figure 7.7. The structure is constructed of layers. The pseudo-Watson-Crick bonding motifs (green for neutral cytosine and red for cytosinium, Figure 7.7) stack on top of each other and the isonicotinate molecules (purple) stack on top of each other. The two different water molecules (black and blue) combine to link the layers together. This gives a three dimensionally hydrogen bonded crystal structure.

7.1.2 1:1:1 Cytosinium Nicotinate Monohydrate

Cytosine and nicotinic acid were dissolved in methanol in a molar ratio of 1:1 and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.3mm by 0.4mm by 0.4mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 7.6.

This molecular complex forms in a 1:1 ratio of cytosine to nicotinic acid and shows 100% proton transfer from the carboxylic acid group of the nicotinic acid molecule to the cytosine. The predominant motif observed is therefore the hydrogen bonded ring motif creating a heterodimer (Figure 7.8). The only base-pairing that would be possible is the type B pseudo Hoogsteen motif; however in this molecular complex, this is not observed.

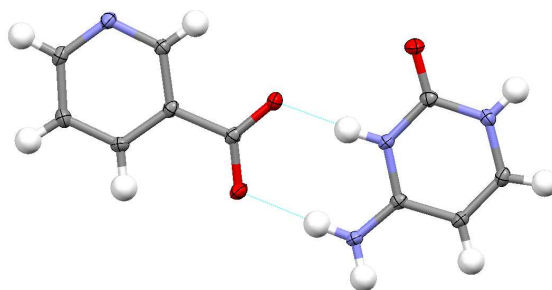


Figure 7.8 Heterodimer with a hydrogen-bonded ring motif in the molecular complex of cytosinium nicotinate monohydrate

The hydrogen-bonded ring is constructed from two N-H...O hydrogen bonds. The first hydrogen bond is from one of the hydrogen atoms of the amine group of the cytosinium to one of the oxygen atoms of the carboxylate group of the nicotinate. This hydrogen bond is of moderate strength and has N...O distance of 2.816(1) Å. The other hydrogen bond that completes the hydrogen bonded ring is from the hydrogen of a protonated ring nitrogen of the cytosinium molecule to the other oxygen atom of the carboxylate group. This bond is also of moderate strength with N...O distance of 2.651(1) Å.

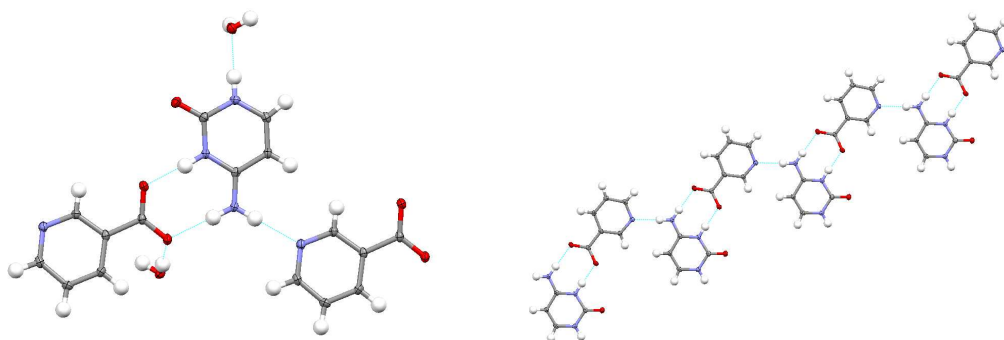


Figure 7.9 Left, hydrogen bonds the heterodimer and right, chains of heterodimers linked by single hydrogen bonds in the cytosinium nicotinate monohydrate molecular complex.

There are a number of hydrogen bonds involving the heterodimers (Figure 7.9, left). The first of these is an N-H...N hydrogen bond from the second hydrogen of the amine group to the ring nitrogen of the nicotinate molecule. This hydrogen bond is of N...N distance 2.899(2) Å and is of moderate strength. This nicotinate molecule lies in the same plane as the heterodimer and the net result is the formation of chains of heterodimers (Figure 7.9, right). The water molecule again performs two roles, both connecting heterodimer chains within layers and connecting parallel layers. The water molecule oxygen atom acts as an N-H...O hydrogen bond acceptor for a hydrogen bond from a protonated ring nitrogen of the cytosine molecule. This hydrogen bond is of N...O distance 2.834(2) Å and is of moderate strength. This water molecule then acts as a linking unit to neighbouring chains of heterodimers. A DHAA bifurcated O-H...O hydrogen bond is formed to a single

heterodimer bridging the hydrogen bonded ring with the major hydrogen bond forming to the carbonyl oxygen atom of the cytosinium molecule and the minor component to one of the carboxyl oxygens of the nicotinate molecule (Figure 7.10, green dotted lines). The major and minor components have O...O distances of 3.078(3) and 3.187(3) Å, respectively. The water molecule takes part in one more hydrogen bond and it involves the hydrogen not involved in the bifurcated hydrogen bonds. This hydrogen bond is between the hydrogen of the water molecule and one of the oxygen atoms of the carboxylate group of a nicotinate in a parallel plane. This hydrogen bond is of moderate strength and is of O...O distance 2.806(2) Å (Figure 7.11). A further weak C-H...O hydrogen bond is formed between cytosinium molecules in neighbouring chains with a C...O distance of 3.082(2) Å (Figure 7.10, violet dotted line).

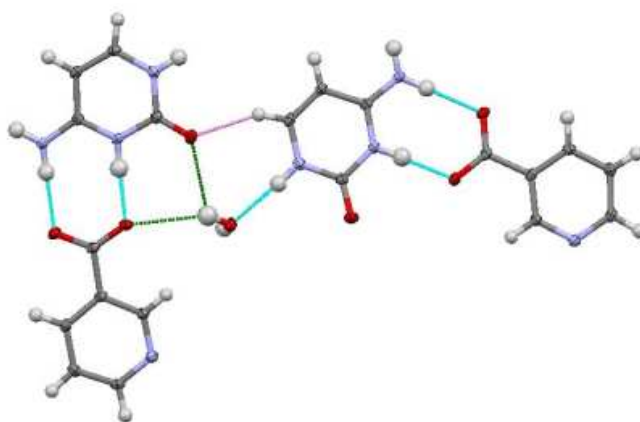


Figure 7.10 The connection of neighbouring heterodimer chains in the molecular complex of cytosinium nicotinate monohydrate. The bifurcated hydrogen bond involving the water molecule is represented by green dotted lines and the weak C-H...O hydrogen bond by a violet dotted line.

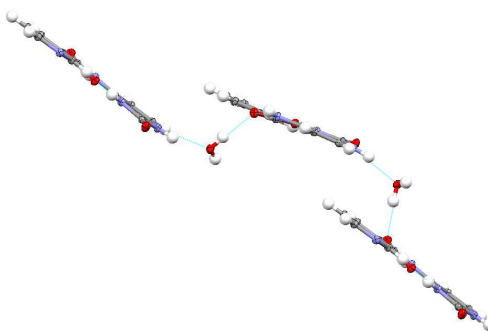


Figure 7.11 Water molecule connecting the layers in the molecular complex of cytosinium nicotinate monohydrate

An eight membered hydrogen bonded ring is formed, constructed of two heterodimer pairs and four individual molecules between two layers (Figure 7.12). Two pairs are connected via the two hydrogen bonds that involve the water molecule and the other three molecules are connected in the ring via single hydrogen bonds.

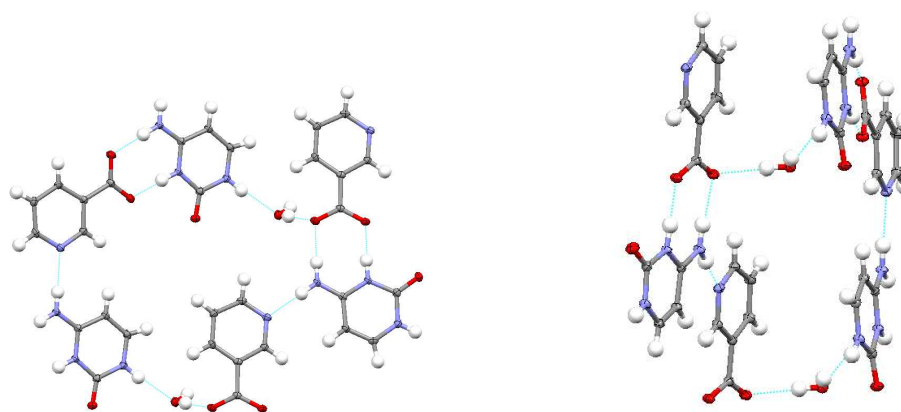


Figure 7.12. Eight molecule hydrogen bonded ring in the molecular complex of cytosinium nicotinate monohydrate

There are also stacking interactions in the form of $\pi\cdots\pi$ stacking interactions (Figure 7.13). Cytosinium molecules are stacked on top of nicotinate molecules, through two stacking interactions. The first of these is present between two layers that are linked together via the hydrogen bonds involving the water molecules. This stacking distance is approximately 3.358 Å. The other stacking interaction is between layers that have no hydrogen bonds between the layers present in the structures, with distance of approximately 3.268 Å.

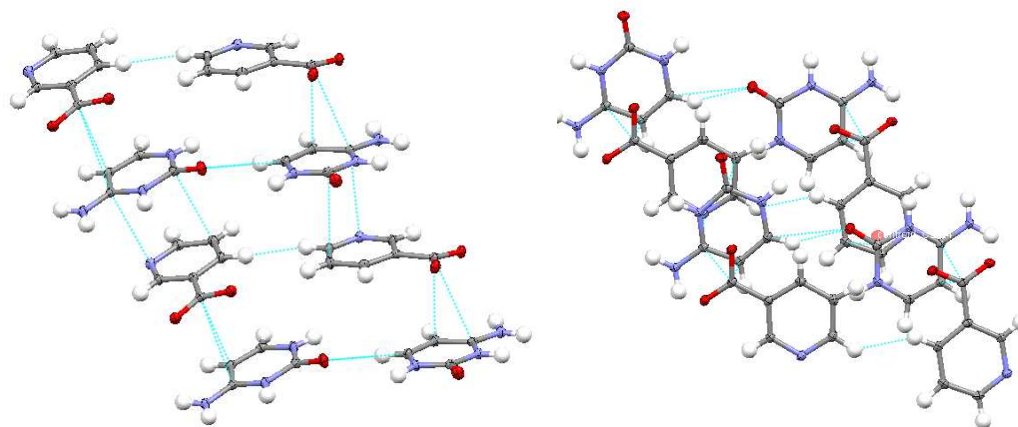


Figure 7.13 (a) Stacking interactions and (b) staggered stacking in the molecular complex of cytosinium nicotinate monohydrate

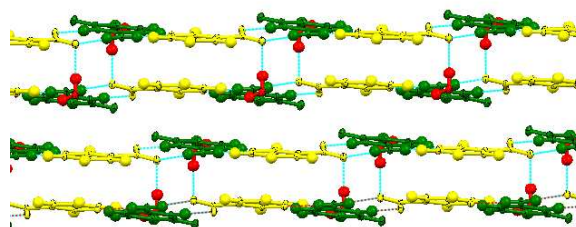


Figure 7.14 Colour coded structure to show planar nature and the role of the water molecule connecting separate layers in the molecular complex of cytosinium nicotinate monohydrate. Water (red), cytosine (green) and nicotinic acid (yellow)

The cytosine (green) and the nicotinic acid (yellow) construct layers lying along the *b* axis. The water molecules (red) combine to link separate chains together and stacking interactions are also present between the chains in addition to the hydrogen bonds

involving water molecules. The water molecules only connect pairs of layers together; there are no hydrogen bonds in the central layer spacing shown in Figure 7.14.

7.2. Molecular Complexes of Cytosine with Substituted Benzoic Acids

7.2.1 2:1 Cytosine and 2-nitrobenzoic acid

Cytosine and 2-nitrobenzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.3mm x 0.3mm x 0.2mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. A summary of the crystallographic data are given in Table 7.6.

The molecular complex forms in a 2:1 ratio of cytosine to 2-nitrobenzoic acid; a proton is transferred from the carboxyl group of the 2-nitrobenzoic acid molecule to the unprotonated heteroatom of one of the cytosine rings. Therefore three species are contained within the asymmetric unit: a protonated cytosinium molecule, a deprotonated 2-nitrobenzoic acid molecule (2-nitrobenzoate) and a neutral cytosine molecule. This means that the pseudo-Watson-Crick base pair hydrogen bonding motif predominately exhibited in similar 2:1 complexes is observed (Figure 7.15, left). The three hydrogen bonds in this unit are all of moderate strength (Table 7.2).

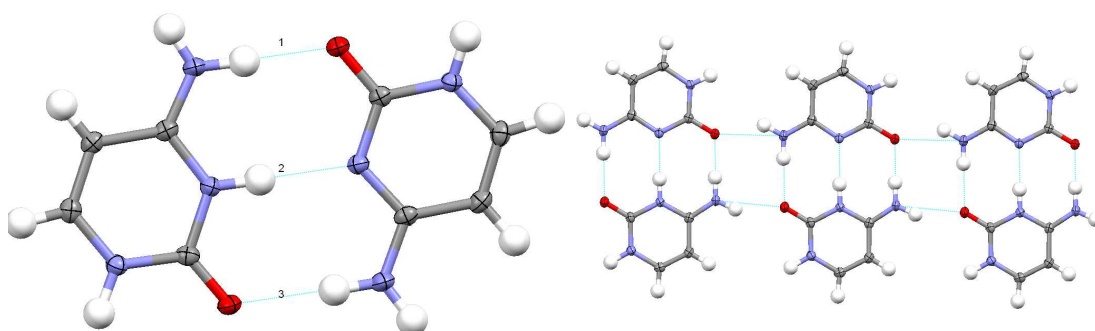


Figure 7.15. Left, pseudo-Watson-Crick base pair unit and right, chains of cytosine molecules in the molecular complex of cytosine and 2-nitrobenzoic acid

Table 7.2 Hydrogen bonds for the pseudo-Watson-Crick hydrogen bond motif in the molecular complex of cytosine and 2-nitrobenzoic acid

Hydrogen Bond	D...A Distance
O...H-N	2.739(2) Å
N...H-N	2.839(2) Å
N-H...O	2.939(2) Å

Chains of cytosine molecules are formed through two independent N-H...O single hydrogen bonds to neighbouring pseudo-Watson-Crick units (Figure 7.15, right). The base paired chain forms in a way that on one side of the chain there are only protonated cytosine molecules and on the other side there are only neutral cytosine molecules. There is a hydrogen bond between two separate neutral cytosine molecules in the chain and one hydrogen bond between two protonated cytosine molecules. The first hydrogen bond is from the amine group of a neutral cytosine to the oxygen atom of a nearby neutral cytosine molecule. This hydrogen bond is of moderate strength and is of N...O distance 2.835(2) Å. The second hydrogen bond is the equivalent between the amine group and the oxygen of protonated cytosine molecules. This N...O hydrogen bond length is 2.832(2) Å. The strength of this hydrogen bond is therefore not affected by the charged state of the cytosine molecule.

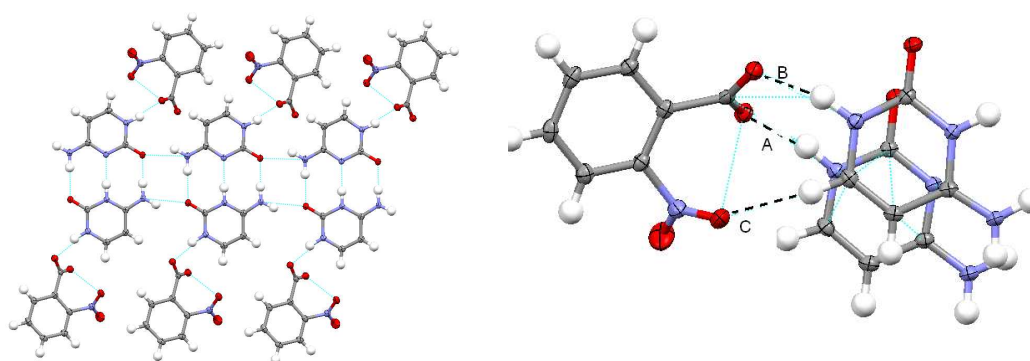


Figure 7.16 2-nitrobenzoate molecules hydrogen bonded to cytosine base-paired chain in the 2:1 cytosine and 2-nitrobenzoic acid molecular complex.

On either side of the cytosine chain the 2-nitrobenzoate molecule is tied to the chain via two similar moderate hydrogen bonds N-H...O hydrogen bonds from a protonated ring nitrogen of the cytosine to the carboxylate group of the 2-nitrobenzoate molecule (Figure 7.16, left). The hydrogen bond involving the neutral cytosine molecule has an N...O distance of 2.742(2) Å (Figure 7.16, right; hydrogen bond A) and the hydrogen bond with the cytosinium molecule has an N...O distance of 2.699(2) Å (hydrogen bond B). There is an additional weak C-H...O hydrogen bond from an aromatic hydrogen to one of the oxygen atoms of the nitro group (hydrogen bond C). This is of C...O distance 3.303 (3) Å.

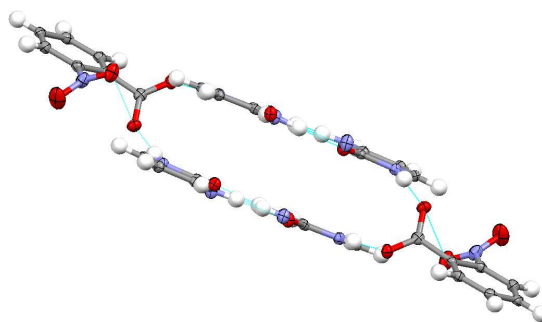


Figure 7.17 2-nitrobenzoate molecules linking parallel cytosine base-paired chains via single hydrogen bonds in the 2:1 cytosine and 2-nitrobenzoic acid molecular complex.

The 2-nitrobenzoate molecules do not lie in the same plane as the cytosine chains. In addition, the carboxylate group is twisted out of the plane of the benzene ring of the 2-nitrobenzoate molecule itself. This enables it to play a connecting role between parallel chains of cytosine through hydrogen bonds A and B creating a six molecule unit (Figure 7.17). The cytosine chains are stacked such that they are inverted relative to one another and thus the neutral part of the chain stacks above the protonated part of the chain below and vice versa.

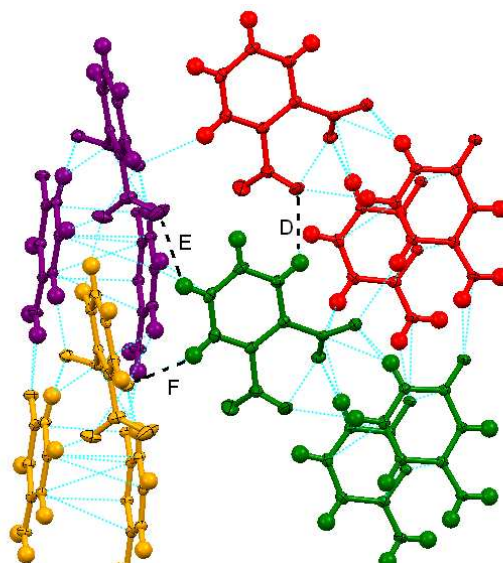


Figure 7.18 Linking between chains using weak hydrogen bonds and the subsequent orientation of the chains involved in the structure of cytosine and 2-nitrobenzoic acid

There are three weak hydrogen bonds involving aromatic hydrogen atoms of the 2-nitrobenzoate molecule. The first involves the hydrogen positioned next to the carboxylate group of the 2-nitrobenzoate and one oxygen atom of a nearby nitro group (Bond D, Figure 7.18). This bond is of moderate strength and is of C...O length 3.291(2) Å. The next involves the hydrogen in the 4- position of the 2-nitrobenzoate and the other oxygen atom of the nitro group of a separate 2-nitrobenzoate molecule. This bond is of weak strength and is of C...O length 3.333(3) Å (Bond E). The final weak hydrogen bond is between the hydrogen in the 3- position of the 2-nitrobenzoate molecule and one of the oxygen atoms

of a carboxylate group of a separate 2-nitrobenzoate molecule. This bond weak with C...O length of 3.469(2) Å (Bond F). These hydrogen bonds combine to create a zig zagged structure where the orientation of the 2-nitrobenzoate molecules and the subsequent interactions produce the change in orientation of nearby chains (Figure 7.19). Bond D connects the red and green units with cytosine chains in parallel layers but staggered with respect to one another. Bonds E and F connect units together which are almost perpendicular to one another.

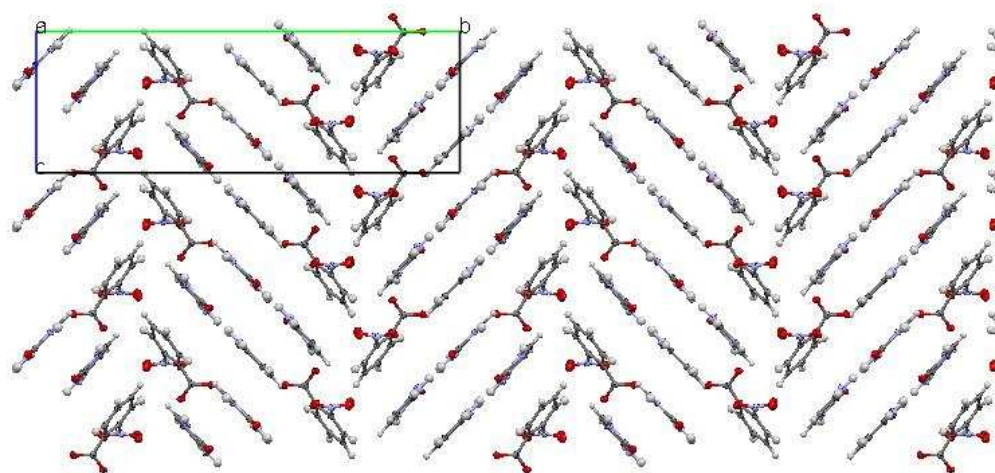


Figure 7.19 The orientation of the building blocks described showing non- planar packing in the structure of cytosine and 2-nitrobenzoic acid

There are some base stacking interactions present in this structure also between the two cytosine chains in one building block (Figure 7.19). These interactions have the pseudo-Watson-Crick base paired chains stacked on top of each other with the charged side of the chain being on top of a neutral side of the chain and vice versa. The stacking distance is approximately 3.130 Å. Figure 7.20 shows that there are reasonably well defined regions of 2-nitrobenzoic acid (green) and regions of cytosine (red) when looking along the a-axis. Weak interactions connect neighbouring building blocks and all the stronger interactions, in the form of moderate hydrogen bonds, are involved in the construction of the basic six molecule building block. The other stacking interactions present are between the 2-nitrobenzoic acid and the cytosine chain directly below it. The 2-nitrobenzoic acid molecule stacks on top of the neutral cytosine molecule with this stacking distance measured as approximately 3.086 Å. There is one final stacking interaction present between separate building blocks, comprising a charged cytosine molecule stacked on top of a charged cytosine. This distance is measured to be approximately 3.207 Å.

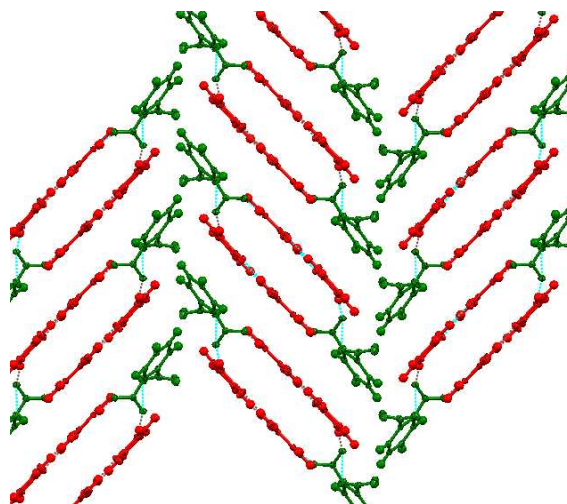


Figure 7.20 Global packing of the cytosine and 2-nitrobenzoic acid molecular complex. Cytosine molecules are shown in red and 2-nitrobenzoate molecules in green.

7.2.2 2:1 Cytosine and 3-bromobenzoic acid

Cytosine and 3-bromobenzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a needle shape and the crystal used for characterisation by X-ray diffraction had approximate dimensions of 0.4mm x 0.2mm x 0.2mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 7.6.

This molecular complex forms in a 2:1 ratio of cytosine to 3-bromobenzoic acid; a hydrogen atom is again transferred from the carboxyl group of the 3-bromobenzoic acid molecule to the unprotonated heteroatom of the cytosine ring. The asymmetric unit therefore contains a protonated cytosine molecule (cytosinium), a deprotonated 3-bromobenzoic acid (3-bromobenzoate) and a neutral cytosine molecule. The presence of both a neutral and a protonated cytosine again results in the formation of pseudo-Watson-Crick base paired hydrogen bonding motifs (Figure 7.21; Table 7.3).

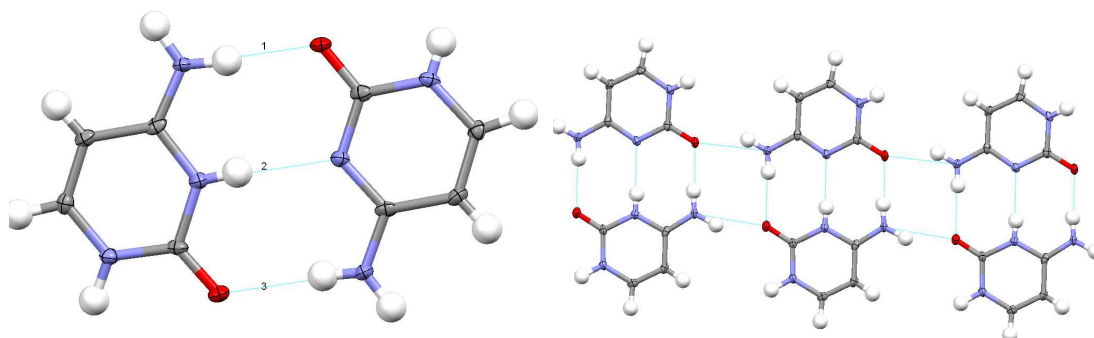


Figure 7.21.Left, pseudo-Watson-Crick base paired hydrogen bonding motif and right, cytosine chains in the molecular complex of cytosine and 3-bromobenzoic acid

Table 7.3. The three hydrogen bonds in the pseudo-Watson-Crick base-pairing motif in the molecular complex of cytosine and 3-bromobenzoic acid

Hydrogen Bond	D...A distance
O...H-N	2.765(4) Å
N...H-N	2.853(4) Å
N-H...O	2.923(4) Å

Cytosine chains are again constructed of pseudo-Watson-Crick base pairs joined through two N-H...O hydrogen bonds (Figure 7.21). The chains have the protonated cytosine molecules on one side and the neutral cytosine molecules on the other side of the chain. There is one hydrogen bond between adjacent neutral cytosine molecules from the amine group of one cytosine to the oxygen atom of another cytosine molecule. This hydrogen bond is of moderate strength and is of N...O distance 2.818(4) Å. The other hydrogen bond is the equivalent between two cytosinium molecules. This hydrogen bond is also of moderate strength and is of N...O distance 2.819(4) Å.

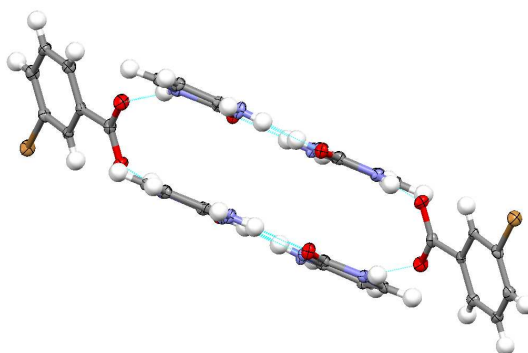


Figure 7.22 3-bromobenzoate molecules linking parallel cytosine chains via single hydrogen bonds in the molecular complex of cytosine and 3-bromobenzoic acid

The 3-bromobenzoate molecules are planar and do not lie in the same plane as the cytosine chains. Two chains of cytosine lie parallel to one another and 3-bromobenzoic acid molecules link them on either side (Figure 7.22). The chains are inverted relative to each other such that neutral side of the chain is above the protonated side of the other chain and vice versa. The 3-bromobenzoate molecules act as hydrogen bond acceptors for two moderate strength N-H...O hydrogen bonds. One of the hydrogen bonds is from the hydrogen on a protonated ring nitrogen of a neutral cytosine to one of the oxygen atoms of a nearby carboxylate group. This hydrogen bond is of moderate strength and is of N...O distance 2.726(4) Å. The other hydrogen bond is the equivalent from the protonated ring nitrogen of the cytosinium molecule to the other oxygen atom of the carboxylate group. This hydrogen bond is also of moderate strength and is of N...O distance 2.611(4) Å.

This structure has some similarities to that of the cytosine 2-nitrobenzoic acid molecular complex (Section 7.2) where weak interactions link nearby six molecule building blocks together. There are two weak interactions, a Br \cdots H-C interaction between the hydrogen next to the normally protonated nitrogen of a cytosine ring and a weak C-H \cdots π interaction between two 3-bromobenzoate molecules. The first of these has a Br \cdots H distance of 3.039 Å which is just within the sum of the van der Waals radii of the two atoms (3.05 Å). The C \cdots π distance is approximately 3.667 Å. It is through these interactions that the chains are orientated in a zig-zag pattern. Both this structure and that of the cytosine 2-nitrobenzoic acid molecular complex have well defined areas of cytosine chains (red, Figure 7.23)) and regions of benzoic acids (green, Figure 7.23)). The similarity is striking considering the increased potential hydrogen bonding in the 2-nitrobenzoic acid complex due to the presence of the nitro group. The six molecule hydrogen bonded unit and the extended cytosine chains persist in both structures.

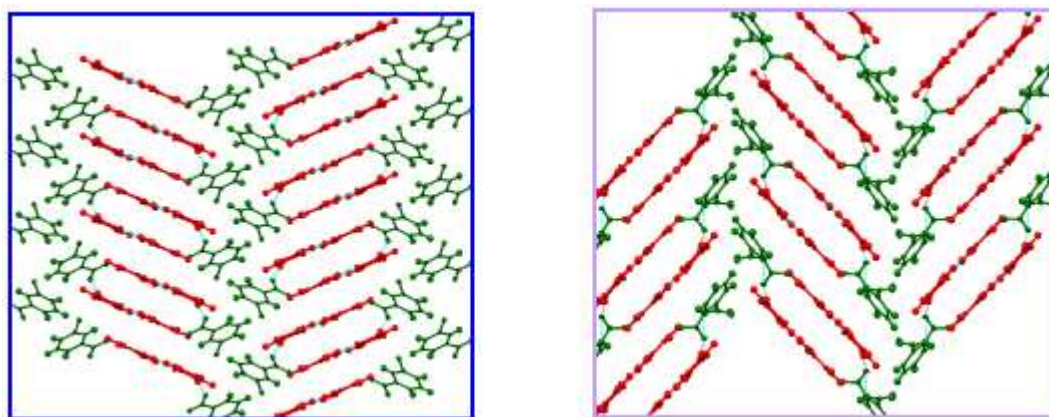


Figure 7.23 Comparison of structures of (a) 3-bromobenzoic acid and cytosine and (b) cytosine and 2-nitrobenzoic acid with cytosine in each structure being represented by red and benzoic acids represented by green.

Base stacking interactions are also present in this bromobenzoic acid complex structure in the same manner as that seen in the nitrobenzoic acid structure. The stacking interactions once again have the neutral cytosine above a charged cytosine with the opposite on the other side of the same chain. This structure has an approximate stacking distance of 3.127 Å. There is also a further stacking interaction between the two charged cytosine molecules between the separate building blocks, of approximate distance 3.119 Å.

7.2.3 1:1:1 Cytosinium 2,4,6-trimethylbenzoate monohydrate

Cytosine and 2,4,6-trimethylbenzoic acid were dissolved in methanol in a 1:1 molar ratio and the solvent was allowed to evaporate slowly until crystals formed. The crystals were colourless with a block shape and the crystal used for characterisation by X-ray diffraction

had approximate dimensions of 0.3mm x 0.3mm x 0.2mm. Data were collected on a Bruker Apex II CCD diffractometer equipped with an Oxford Cryosystems Helix at 100K. Crystallographic data are summarised in Table 7.6.

The molecular complex forms in a 1:1 ratio of cytosine to 2,4,6-trimethylbenzoic acid and proton transfer is found from the carboxyl group of the trimethylbenzoic acid molecule to the unprotonated heteroatom of the cytosine ring. There is therefore one protonated cytosine molecule (cytosinium), one deprotonated 2,4,6-trimethylbenzoic acid molecule (2,4,6-trimethylbenzoate) and one water molecule in the asymmetric unit.

A heterodimer between the cytosinium and the 2,4,6-trimethylbenzoate molecules is formed, further supporting the observation that this hydrogen bonding motif is predominant in molecular complexes showing proton transfer and a 1:1 ratio of the cytosine to the co-molecule (Figure 7.24). The hydrogen bonded ring of the heterodimer is constructed from two N-H...O hydrogen bonds, the first of which is from a hydrogen on the amine group of the cytosinium to one of the oxygen atoms on the carboxylate group of the 2,4,6-trimethylbenzoate. The second hydrogen bond is from the hydrogen on a protonated ring nitrogen of the cytosinium to the other oxygen of the carboxylate group. These hydrogen bonds are both of moderate strength and have N...O distances of 2.758(1) Å and 2.772(1) Å, respectively.

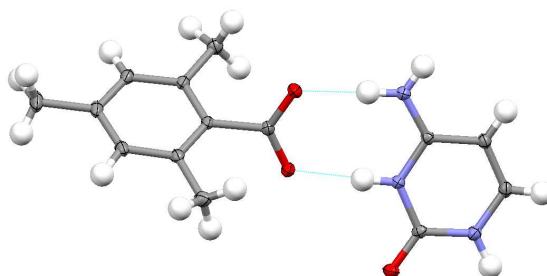


Figure 7.24 Hydrogen-bonded ring motif in the cytosinium 2,4,6-trimethylbenzoate monohydrate molecular complex.

Whilst the pseudo-Watson-Crick base pair motif is not possible in this complex, the type B pseudo-Hoogsten base pair motif is possible. However, this is not found here, with no extended homo-base pairing chain found in this structure. There is however, a chain of cytosinium molecules connected by a single N-H...O hydrogen bond from the amine group one cytosinium to the carbonyl oxygen on a second (Figure 7.25, left). These hydrogen bonds are of moderate strength and of N...O distances of 2.882(1) Å.

The carboxylate group of the 2,4,6-trimethylbenzoate molecule acts as a hydrogen bond acceptor for two hydrogen bonds involving water molecules (Figure 7.25, right). The two

hydrogen bonds are of moderate strength with O...O distances of 2.801(1) Å (A, Figure 7.25, right) and 2.760(1) Å (B, Figure 7.25, right). The first of these hydrogen bonds acts as an additional linker molecule between cytosinium molecules in the hydrogen bonded chain (Figure 7.25, right). An N-H...O hydrogen bond is formed, between a protonated ring nitrogen of the cytosinium and the oxygen atom of the water molecule. This bond is of moderate strength and is of N...O distance 2.739(1) Å. This creates a four molecule $R_4^3(10)$ hydrogen bonded ring. The second water hydrogen bond connects cytosinium chains that lie approximately perpendicular to one another (Figure 7.26, left). These almost perpendicular units are also held together through C-H...O hydrogen bonds from a methyl hydrogen atom to the carboxylate of a 2,4,6-trimethylbenzoate molecule creating a twisted chain of these molecules (Figure 7.26, right). These hydrogen bonds are weak with a C...O distance of 3.556 (2) Å.

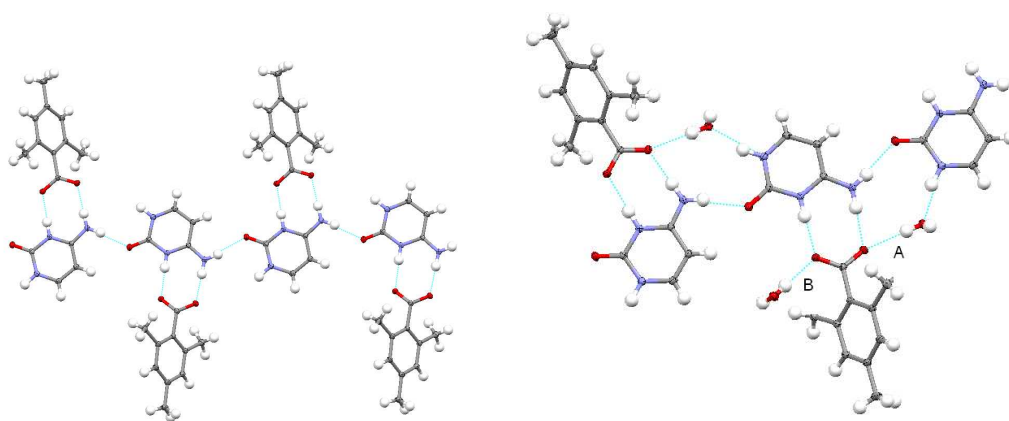


Figure 7.25 Left, chain of cytosinium molecules linked by single hydrogen bonds and right, hydrogen bonds involving the water molecules in the molecular complex of cytosinium 2,4,6-trimethylbenzoate monohydrate

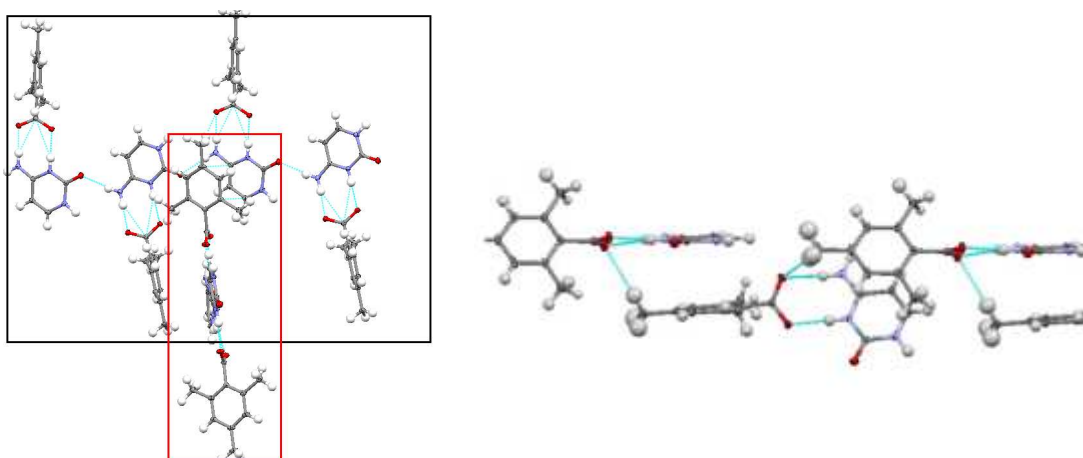


Figure 7.26 Orientation (approximately perpendicular) of neighbouring cytosinium chains in the molecular complex of cytosinium 2,4,6-trimethylbenzoate monohydrate

The base stacking in this structure can be clearly seen in Figure 7.27. There are no cytosinium cytosinium stacking interactions as the 2,4,6-trimethylbenzoate molecules are

stacked on top of the cytosinium molecules. The distance between the two layers is approximately 3.289 Å.

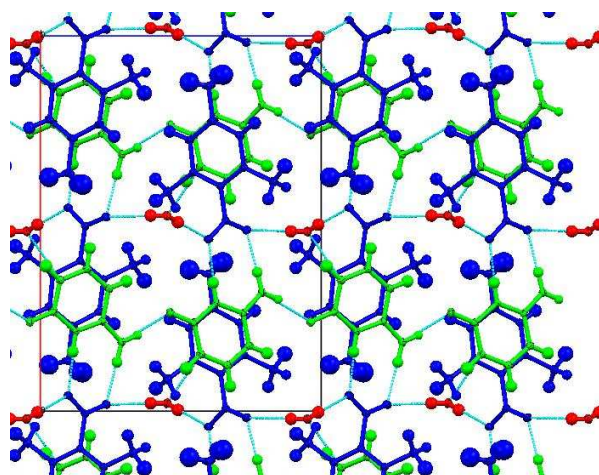


Figure 7.27 Stacking of cytosinium on top of 2,4,6-trimethylbenzoate in the molecular complex of cytosinium 2,4,6-trimethylbenzoate monohydrate

7.3 Conclusions

Where the crystallisation ratio is 2:1 cytosine to co-molecule, at least one neutral and one charged molecule are found in the resulting structure. All of the 2:1 structures reported in this chapter contain a cytosine chain constructed of pseudo-Watson-Crick base-pairs with each base pair formed from one cytosine and one cytosinium molecule. These base pairs are connected to each other through single hydrogen bonds forming an extended chain. There is base stacking present in all of the structures with chains of cytosine sitting directly on top of each other with the charged cytosinium on top of the non-charged cytosine molecule. The non-cytosine molecules in the structures (co-molecules or water molecules) link these chains together. Where the crystallisation ratio in the molecular complex is 1:1, a fully ionic molecular complex tends to be found where the primary hydrogen bonding is the formation of a heterodimer held together by a hydrogen bonded ring. Many of these structures are hydrates and the presence of water molecules does not interrupt the predominant hydrogen bonding motifs.

The ΔpK_a (Table 7.4, Table 7.5) values indicate that hydrogen transfer might be expected in all of these complexes. The complex most likely to show proton transfer with the largest ΔpK_a value is that of the complex with nicotinic acid. However this, perhaps surprisingly, forms in a 2:1 ratio of cytosine to isonicotinic acid. The pK_a values are thus useful as a guide to whether or not proton transfer will occur, but it has only limited validity in predicting the crystallisation ratio and therefore the degree of relative hydrogen transfer. Also the base stacking present in the structures will predominantly have charged cytosine molecules stacked on top of neutral cytosine molecules and the only point where

there are stacking interactions between the two charged molecules is between the building blocks; this may be a consequence of the preferred bonding pattern adopted. However, when there are only charged cytosinium molecules present these are also seen to avoid stacking on top of one another and will be present with the cytosinium on top of the co-molecule present in the structure.

Table 7.4 Ratio of products, percentage of cytosine molecules gaining a hydrogen and ΔpK_a values for the structures in Chapter 7

Co-crystal	Crystallisation ratio	Hydrogen Transfer with respect to cytosine	ΔpK_a
Cytosine and 2-nitrobenzoic acid	2:1	50%	2.44
Cytosine and 3-bromobenzoic acid	2:1	50%	0.76
Cytosine and Isonicotinic acid Hydrate	2:1:2	50%	2.83
Cytosinium 2,4,6-trimethylbenzoate monohydrate	1:1:1	100%	1.17
Cytosinium nicotinate monohydrate	1:1:1	100%	2.60

Table 7.5 pK_a values of the starting materials

Starting Material	pK_a Value
Cytosine	4.60
2-nitrobenzoic acid	2.16
3-bromobenzoic acid	3.81
Isonicotinic acid	1.77
2,4,6-trimethylbenzoic acid	3.43
Nicotinic acid	2.00

Table 7.6. Table of crystallographic data for the molecular complexes of cytosine in Chapter 7.

Compound	Cytosine and isonicotinic acid hydrate	Cytosinium nicotinate monohydrate	Cytosine and 2-nitrobenzoic acid	Cytosine and 3-bromobenzoic acid	Cytosineium 2,4,6-trimethylbenzoate monohydrate
Formula	C14 H19 N7 O6	C10 H12 N3 O4	C15 H15 N7 O6	C15 H15 N3 O4 Br1	C14 H19 N3 O4
Crystallisation Conditions	Methanol, Room Temperature	Methanol, Room Temperature	Methanol, Room Temperature	Methanol, Room Temperature	Methanol, Room Temperature
Molecular weight / gmol⁻¹	381.34	238.22	389.32	381.20	293.32
Temperature (K)	100 K	100 K	100 K	100 K	100 K
Space Group	P-1	P2 ₁ /n	P2 ₁ /c	P2 ₁ /c	Pna2 ₁
<i>a</i> (Å)	7.0662(8)	7.4120(6)	7.3848(1)	7.3543(1)	16.0704(14)
<i>b</i> (Å)	7.3657(7)	18.2353(14)	26.0769(3)	32.4834(6)	7.4773(7)
<i>c</i> (Å)	17.6043(18)	8.3167(6)	9.1863(1)	7.3020(1)	12.0458(10)
α (°)	100.776(7)	90	90	90	90
β (°)	92.245(7)	103.109(4)	108.556(1)	109.947(1)	90
γ (°)	107.966(7)	90	90	90	90
Volume (Å³)	851.59(27)	1094.79(11)	1511.52(3)	1639.75(6)	1447.46(2)
Z	2	4	4	4	4
θ range (°)	1.2-29.3	2.2-28.7	1.6-24.7	1.3-27.5	2.5-33.5
Completeness	99.20%	95.90%	99.90%	99.90%	97.90%
Reflections Collected	44572	36945	26556	25578	81430
Independent	4614	2648	2869	3752	5448
Refln (obs.I>2theta(I))	3457	2151	2391	2641	5190
R_{int}	0.0623	0.0490	0.0418	0.1021	0.0472
Parameters	320	211	313	295	266
GooF on F²	1.028	1.044	1.034	1.050	1.107
R₁ (Observed)	0.041	0.040	0.030	0.048	0.037
R₁ (all)	0.061	0.055	0.041	0.088	0.041
wR₂ (all)	0.116	0.114	0.079	0.130	0.079

Chapter 8. Systematic Study of Hydrogen Bonding Base Pairing Motifs using DSNAP and the CSD

8.1 Introduction to dSNAP

The dSNAP program has become a very useful tool for comparing related structures in a quick, easy and unbiased manner. This program was developed to help with the volume of crystal structures that are now in the CSD [137], with over 500000 entries now present in that database. The core idea behind dSNAP is that similar fragments defined by their geometry are grouped into clusters using statistical methods. The target molecules or defined regions are selected and the regions defined are decided by the user to give a narrow or wide search. The results can be visualised as a dendrogram and an MMDS (metric multi-dimensional scaling) plot. DSNAP does not provide any direct information on why each cluster is formed, this interpretation being left to the structural chemist carrying out the calculation.

The dSNAP program has been used to analyse both molecular geometry and intermolecular interactions. The work carried out by Parkin et al in 2005 [138] helped to illustrate this through the analysis of the interactions of carboxylic acids with primary amides and with other carboxylic acid groups. In a follow-up paper Collins et al in 2006 [139] analysed interactions of carboxylic acids with secondary amides. The searches were conducted using the CSD and the returned hits were entered into the dSNAP program for analysis. The work allowed the intermolecular interactions present in related structures in the CSD to be examined.

The main conclusions drawn from these initial studies were that the dSNAP program could be used to help analyse hydrogen bonding motifs in both large and small datasets that have been taken from the CSD. This allows for information to be gathered and analysed in a quick manner that would allow new projects to be undertaken in analysing the CSD that would not be previously possible due to time constraints. One major factor, however, is that the clusters, when returned by dSNAP, must be checked by the user due to dSNAP not including any of the chemistry behind the interactions. The program requires no chemical input, and the analysis is carried out purely statistically, with no chemical bias introduced in the analysis. A dSNAP analysis will help the user to identify the common interactions and bonding motifs that can be undertaken by the structures being analysed. The combination of dendrograms, MMDS plots and other display techniques give the user the

easy task of identifying problems in the analysis or help to confirm and interpret the analysis carried out by the program.

8.2. Base Pairing

There are many crystal structures involving DNA bases in the Cambridge Structural Database (CSD). A survey was thus undertaken in two parts: the first was to utilise the cluster analysis program, dSNAP, to investigate the common base-pair motifs between DNA base pairs and to investigate systematically the relative orientation of the two base pairs; the second was to identify the common base pair motifs already recorded in the CSD involving cytosine and co-molecules, with this search not restricted to other DNA bases.

DNA base pairs are constructed of purines and pyrimidines. In each of the two categories there are two DNA bases, adenine and guanine, and thymine and cytosine, respectively (Figure 8.1).

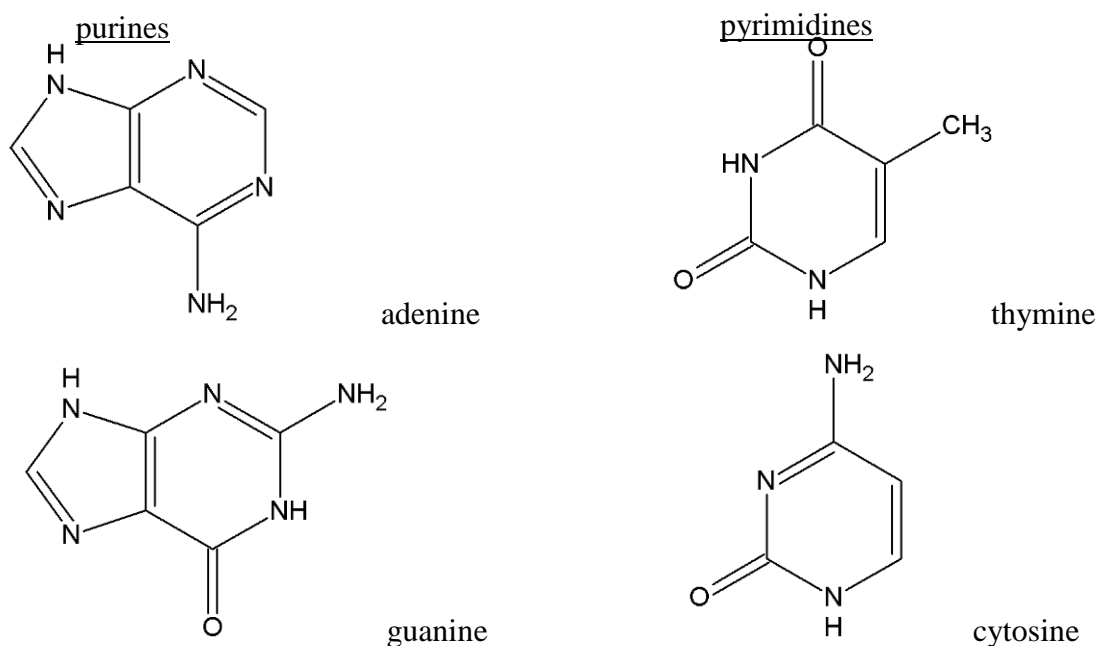


Figure 8.1. Molecular structures of the purine pair, adenine and guanine, and the pyrimidine pair, thymine and cytosine.

The DNA double helix contains two chains of nucleotides, which are linked together through the hydrogen bonding between the nucleic acid bases. When part of the DNA double helical structure, each base specifically pairs to one other base. Adenine forms a hydrogen bonded pair with thymine alone, while guanine and cytosine will form hydrogen bonds exclusively with one another. Altogether, however, there are 29 possible combinations of bases to form both hetero and homo base pairs [140]. Base pairs can be

characterised by a number of parameters including: propeller twist, buckle, C1'-C1' separation and hydrogen bonding parameters [103].

8.2.1 Propeller Twist

The bases of a DNA pair are not precisely coplanar and are twisted with respect to one another. A base pair in DNA is made up of two ringed planar structures joined together by hydrogen bonds. The opposite rotation of these two molecules about their lengthwise axis is referred to as a propeller twist (Figure 8.2).

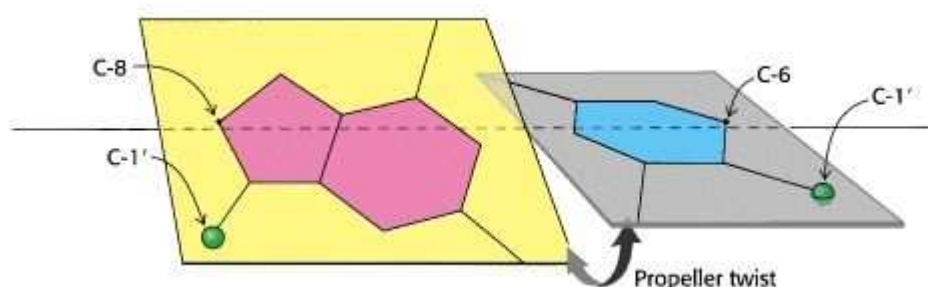


Figure 8.2. Propeller twisting in a DNA base pair [38].

8.2.2 Buckle

This can be classed as the angle between the normals of the planes on which each of the bases lie [103]. Buckling occurs as a result of the levelling of the two bases to eliminate the propeller twist.

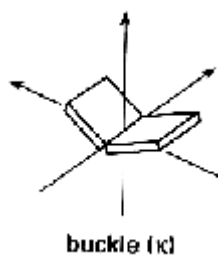


Figure 8.3. A sketch to demonstrate buckle in a base pair motif [39].

8.2.3 Hydrogen Bonding

The last of the classification methods and of most interest in this study is hydrogen bonding. This includes the number and type of hydrogen bonds involved in the base pair, together with the hydrogen bond lengths and angles [141].

8.2.4 dSNAP Study of Base Pairing

An initial search in the CSD showed that base-pairing between adenine and cytosine and guanine and thymine was uncommon, and so the focus was kept to the natural DNA base pair motifs of thymine and adenine, and guanine and cytosine. dSNAP was utilised as a method to cluster the structural fragments and to identify if any trends could be found in the hydrogen bonding motifs, or in the relative orientations of the two hydrogen bonded molecules. dSNAP is an ideal tool for this as the clustering is statistical and thus unbiased.

8.3 dSNAP Study of Adenine and Thymine Base Pairing

There are a large variety of conformations available for the adenine thymine base pair. It is the only pair for which Hoogsteen, Reverse Hoogsteen, Watson Crick and Reverse Watson Crick hydrogen bond types are available. These conformations are illustrated in Figure 8.4.

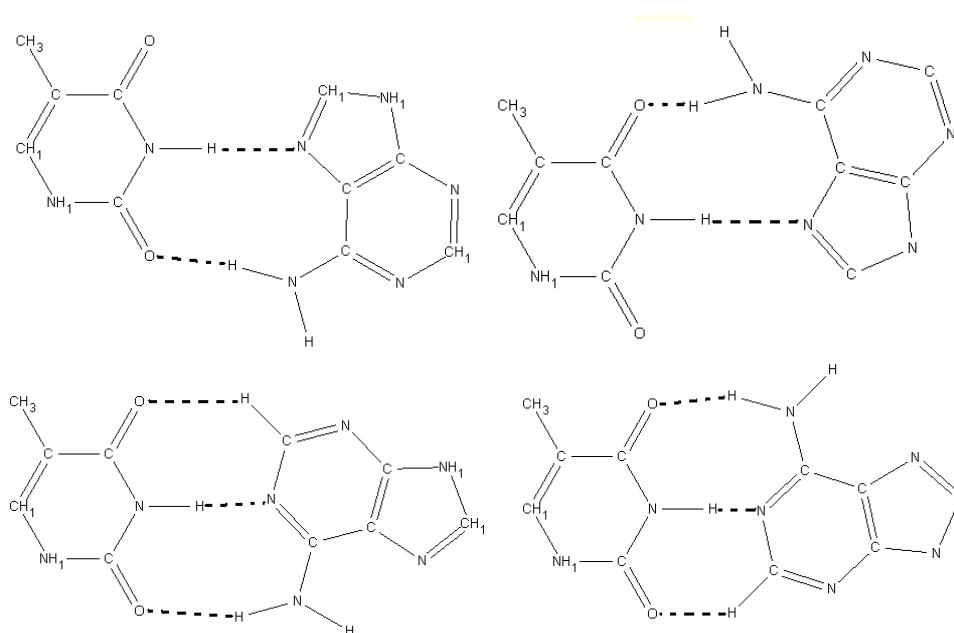


Figure 8.4. The four types of base pairing for the adenine and thymine combination. (a) Reverse Hoogsten (b) Hoogsten (c) Reverse Watson-Crick (d) Watson Crick.

The adenine thymine search was constructed by defining two groups, to represent the base pairing regions between adenine and thymine (Figure 8.5). The intermolecular contacts were defined as being the distances between the closest atoms and not the centroid of the group. The search was restricted to include only structures where the 3D coordinates were reported, there was no disorder, no errors, the structures were not polymeric, there were no ions or powder structures and the structures were only organic. This search returned 18 hits and these were input into dSNAP and the dendrogram shown in Figure 8.6 was obtained.

Eleven clusters were obtained, six of which contained only one hit, and each of these could be explained in terms of the hydrogen-bonding motif.

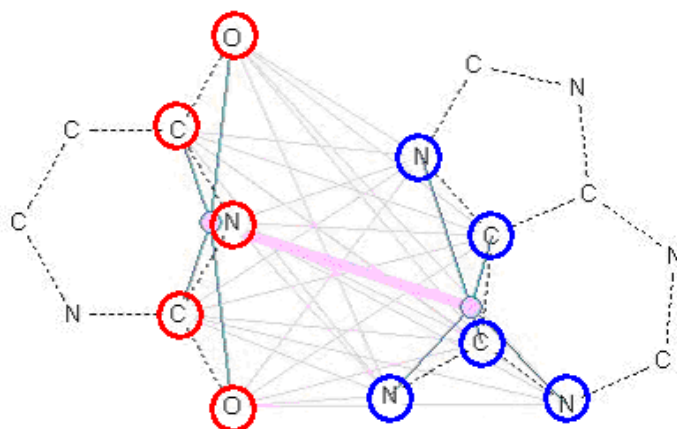


Figure 8.5. The group of atoms on the thymine molecule are circled in red and the group of atoms on the adenine molecules are circled in blue. Close contacts were defined as the distance between nearest atoms.

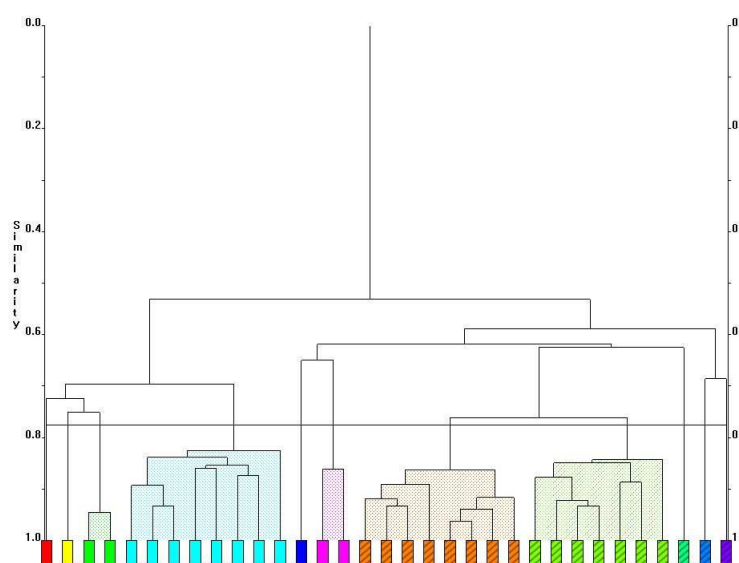


Figure 8.6. Dendrogram obtained from the dSNAP program for the adenine and thymine CSD search.

Another way to visualise and assess the reliability of the clustering is through an MMDS plot. Good clustering is indicated by tightly spaced groupings of each of the hits into their respective clusters, and good separation of the clusters from one another. Figure 8.7 shows the MMDS plot for the clustering of the adenine and thymine complexes and it clearly shows good clustering of the hits. The yellow, green and red hits are very closely related but increasing the cut level to enable them to be clustered together resulted in merging of a large number of other clusters and thus they were kept separate.

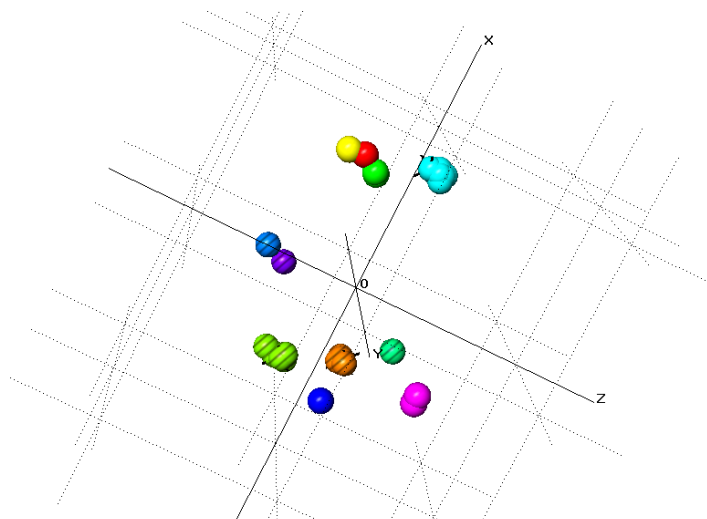


Figure 8.7. MMDS plot of the clustering of adenine and thymine base pairs obtained from the dSNAP program.

8.3.1 Clustering of Adenine Thymine Base Pairs

The clustering of hits using dSNAP effectively separates the fragments into clusters based upon the different types of base pairing. A summary of the clustering is given in Table 8.1.

Table 8.1 – Hydrogen bonding motifs in the dSNAP analysis of adenine thymine base pairs.

Cluster	Colour	Description	No. of Fragments
A	Red	Reverse Hoogsteen	1
B	Yellow	Reverse Hoogsteen	1
C	Green	Reverse Hoogsteen	2
D	Cyan	Hoogsteen	8
E	Blue	Edge to Edge Stacking Interactions	1
F	Pink	Hoogsteen plus single hydrogen bond	2
G	Orange Striped	Watson Crick	8
H	Green Striped	Reverse Watson Crick	7
I	Cyan Striped	N5 to O2 hydrogen bond and reverse Hoogsteen	1
J	Blue Striped	Water and reverse Watson Crick	1
K	Purple Striped	Contains short contacts between plains, reverse Watson Crick and Hoogsteen	1

8.3.1.1 Analysis of the Red, Yellow and Green Clusters

The red and yellow clusters contain only one hit each and the green cluster contains two hits. All three of these clusters are located close to one another in the MMDS plot and correspond to the reverse Hoogsteen hydrogen-bonding motif and thus might be considered to be better described within a single cluster. However, adjusting the cut level resulted in the merging of other clusters, which do not fit well with one another. The cut level was therefore left such that these three hits lie in separate clusters.

The red and yellow clusters are different fragments from the same structure (ADRBFT10 [142]) with two independent reverse Hoogsteen hydrogen-bonding motifs (Figure 8.8). The hydrogen bond distances are significantly different in each of the base pairings (Table 8.2) and thus results in the placement of these units into separate clusters. The dSNAP analysis therefore appears to be sufficiently sensitive to both the orientation of the bases relative to one another and hence the hydrogen bonding motif, but also the strength of the hydrogen bonds. This difference in hydrogen bonds and the classification into two clusters may also be due to the quality of the data as the model for ADRBFT10 has an R-value of 22%. If the quality of the data was improved, then it might be expected that these two units would group into a single cluster.

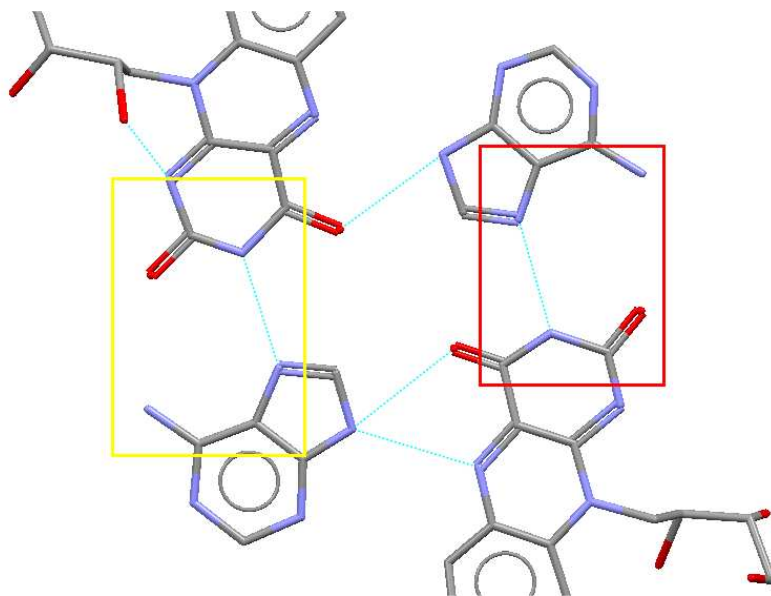


Figure 8.8. Reverse Hoogsteen hydrogen bonding motif in the red and yellow clusters [ADRBFT10] [142].

Table 8.2. Hydrogen bond distances between the base pairs in the red, yellow and green clusters.

Hydrogen Bond	Yellow	Red	Green	
CSD REFCODE	ADRBFT10 [142]	ADRBFT10 [142]	KECYUN [143]	MTYDAP [144]
N...O (Å)	3.126	3.214	2.971	2.985
N...N (Å)	2.792	2.754	2.886	2.878

Both hits in the green cluster also exhibit the reverse Hoogsteen base pairing motif (Figure 8.9). However, the intermolecular hydrogen bond distances in these hits are significantly different than those found in the yellow and red clusters; the two hydrogen bond distances are much more symmetrical. This explains the clustering of the two hits into the green cluster whilst the close grouping of the clusters in the MMDS plot is likely to be a consequence of the similar reverse Hoogsteen hydrogen bonding base pair motif.

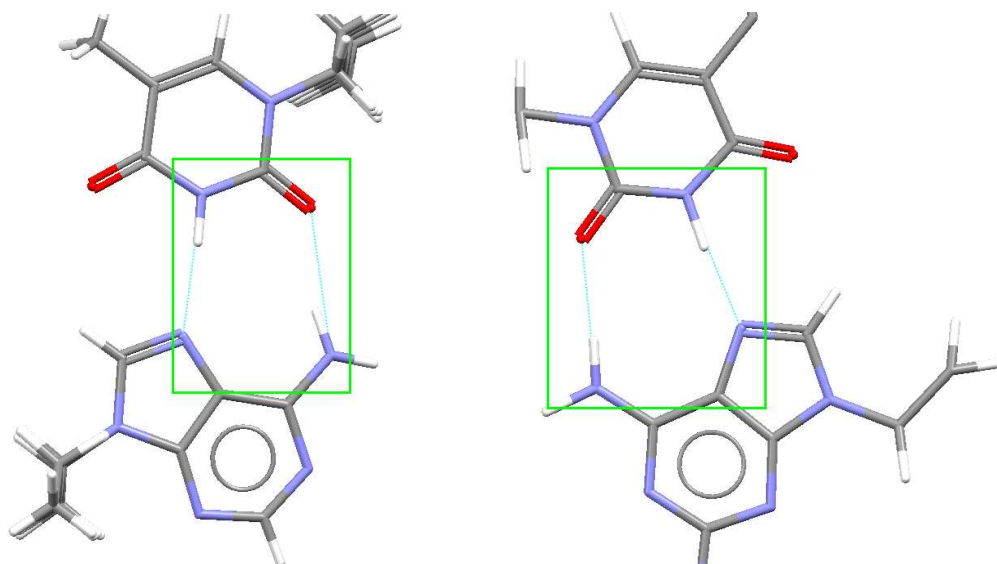


Figure 8.9. The reverse Hoogsteen base pair motifs of the hits in the green cluster. KECYUN (left) and MTYDAP (right).

The base pair in the red cluster contains a slight buckling of the base pair motif with a slight propeller twist in the bonding motif. This was calculated using a plane along the planar ring part of each molecule individually and then comparing the orientation of each plane to calculate the angle between the two molecules. The buckling and propeller twist combine to produce an angle of roughly 5° between the two planes of the base pair. This is a very small effect. The base pair in the yellow cluster contains no propeller twist and only shows a buckling, which is of approximately 5°.

The first hit of the green cluster (KECYUN [143]) does not contain any propeller twist or buckling. The second hit of the green cluster (MTYDAP [144]) does contain a significant propeller twist and no buckling (Figure 8.10). The propeller twist is approximately 7° in

this case. Within the green cluster, the overriding factor for the clustering must be the hydrogen bonding and the clustering is not sensitive to this level of twisting.

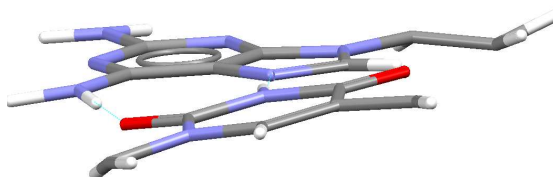


Figure 8.10. Propeller twist in the MTYDAP [144] hit of the green cluster of the adenine thymine dSNAP analysis.

8.3.1.2 Analysis of the Cyan Cluster

The Cyan cluster is the only cluster that contains solely the standard Hoogsteen base-pairing motif and all eight hits contain this motif (Figure 8.11). There is some variation in the hydrogen bonding distances, although in this case it does not affect the clustering (Table 8.3). This may support the premise that the hydrogen bond distances did not have the most significant effect on the clustering for the red and yellow clusters, their separation being more a consequence of the data quality. The clustering again appears to be insensitive to propeller twist and buckling.

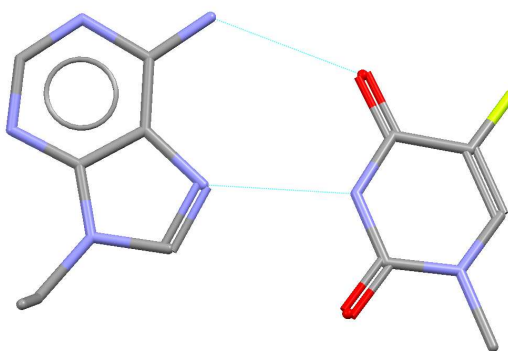


Figure 8.11. Primary hydrogen bonding motif exhibited by all members of the cyan cluster illustrated by ETABFU [148].

Table 8.3. Hydrogen bond distances between the base pairs in the cyan cluster.

Hydrogen Bond	N...O (Å)	N...N (Å)	Buckling or Propeller Twisting
EADBAC [145]	3.125	2.807	Both
EADBAR10 [146]	3.387	2.826	Both
QEDFEK [147]	3.313	2.825	Neither
ETABFU [148]	2.969	2.780	Propeller Twisting
MTHMAD [149]	2.846	2.924	Neither
IMUEAD [150]	3.062	2.820	Buckling
MIUDAP10 [151]	2.979	2.825	Buckling
MBUMAD10 [152]	2.974	2.862	Propeller Twisting

8.3.1.3 Analysis of the blue, green striped and orange clusters

Whilst the blue and green striped clusters both show the reverse Watson-Crick primary hydrogen bonding motif, they are well separated from one another in the dendrogram. The blue cluster has only one hit (BARAAD [153]); however, this molecule is symmetrical and so there are three fragments arising from this unit - the second fragment appears in the green striped cluster and the third in the orange striped (Figure 8.13). The fragment in the blue cluster is separated on its own due to this hit being characterised by distances corresponding to stacking interactions between the two faces of the adenine and thymine molecules (Figure 8.12, left). This is a clear reason for the separation as there are other clusters with stacking interactions with different atoms involved. There are two base pairing motifs in the BARAAD [153] structure, and thus this fragment is placed within the green striped cluster for the reverse Watson-Crick motif (Figure 8.12, right) and also the orange striped cluster with a Watson-Crick motif. The final fragment from this hit is placed in the blue cluster, which represents the stacking interactions that are present in the structure between the defined faces of the starting molecules.

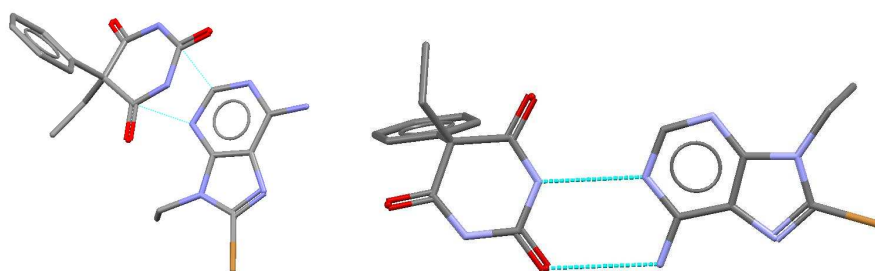


Figure 8.12. BARAAD [153] appearance in the blue and green striped clusters. Left, unit in the blue cluster showing stacking interactions and right, unit in the green striped cluster showing the reverse Watson-Crick motif.

All seven fragments in the green striped cluster show the reverse Watson Crick base-pairing motif. However, not all of the structures show the three hydrogen bonds normally associated with this motif. The only hydrogen bond present in all of the structures is the N-H...N hydrogen bond in the middle of the three. Therefore, the determining factor for this cluster is the presence of the middle hydrogen bond when the molecules are in a position to form a reverse Watson-Crick bonding motif. Once again the effects of the buckling and propeller twisting are too subtle for any real influence on the clustering.

Table 8.4. Hydrogen bond distances for the reverse Watson-Crick primary hydrogen bonding motif of the blue and green striped clusters.

Hydrogen Bond	N...O (Å) (thymine to adenine)	N...N (Å) (thymine to adenine)	N...O (Å) (thymine to adenine)	Buckling or Propeller Twisting
BARAAD_03 [153]	2.965	2.800	N/A	Propeller Twisting
BURBAD [157]	3.056	2.849	N/A	Propeller Twisting
CETQUO [156]	3.090	2.797	N/A	Buckling and Propeller Twisting
FUJDET [157]	2.955	2.859	N/A	Propeller Twisting
IMUEAD [150]	3.156	2.802	N/A	Buckling and Propeller Twisting
MIUDAP10 [151]	2.917	2.909	2.872	Buckling
EBPETY [159]	3.017	2.943	2.868	Buckling

The N...O distances were distinguished by taking the first one as being on the same side as the double bond in the 5 membered ring of the adenine molecule.

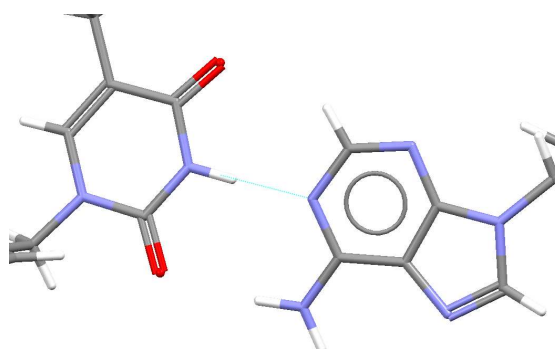


Figure 8.13. Primary hydrogen bonding motif exhibited by all members of the green striped cluster illustrated by CETQUO [156].

8.3.1.3 Analysis of the Pink Cluster

Both the hits within the pink cluster show the Hoogsteen base pair and therefore also appear in the cyan cluster. However there is an additional single hydrogen bond between one of the oxygen molecules of the thymine unit and the amine group of the adenine unit in both (Figure 8.14). This is the only other intermolecular hydrogen bond between the two defined areas so must be the reason for this distinct cluster (Table 8.5). This cluster is not defined by a base pair and thus will not show propeller twisting or buckling.

Table 8.5. Hydrogen bond distance of single hydrogen bond between the adenine and thymine in pink cluster

Bond	ETABFU [148]	MTHMAD [149]
N...O (Å)	2.973	2.881

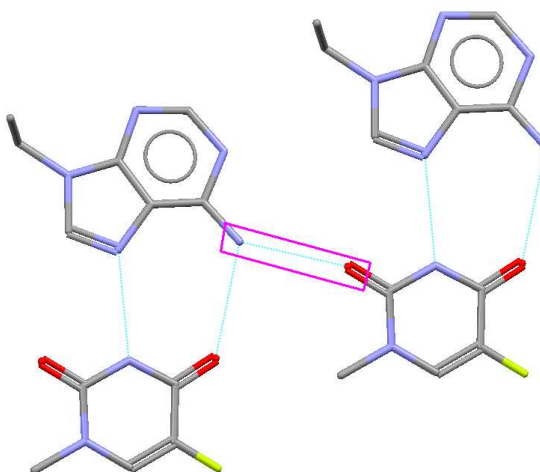


Figure 8.14. Single hydrogen bond between adenine and thymine in the pink cluster illustrated by ETABFU [148].

8.2.1.4 Analysis of Orange Striped Cluster

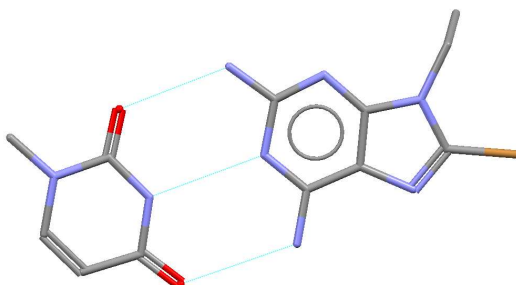


Figure 8.15. Primary bonding motif exhibited by all members of the orange striped cluster illustrated by EBPMUR [155].

Table 8.6. Hydrogen bond distances in the Watson Crick base pair between the adenine and thymine in orange striped cluster.

Hydrogen Bond	O...H-N (Å)	N...H-N (Å)	O...H-N (Å)	Buckling or Propeller Twisting
BARAAD_02 [153]	3.195	2.780	N/A	Propeller Twisting and Buckling
EADBAR10 [146]	3.218	2.832	N/A	Buckling
QEDFEK [147]	3.313	2.825	3.339	Buckling
EADBAC [145]	3.339	2.790	N/A	Buckling and Propeller Twisting
DETVAA_01 [154]	2.924	2.859	N/A	Propeller Twisting
DETVAA_02 [154]	2.919	2.903	N/A	Propeller Twisting
MTYDAP [144]	2.990	2.987	2.888	Buckling and Propeller Twisting
EBPMUR [155]	2.853	2.952	2.977	Neither

All eight fragments in the orange striped cluster form the standard Watson Crick base pairing. These structures are related as they all contain a single hydrogen bond from the pyrimidine base to the purine base from one of the protonated nitrogen atoms in the ring of the pyrimidine base to the non-protonated nitrogen of the next to the amine group of the purine base (Figure 8.15). The bases are positioned correctly for a Watson-Crick bonding motif, however, it becomes difficult to calculate distances for these due to the poor quality of several of the structures and as a result the hydrogen atoms not being defined. Of the eight hits in this cluster, there are only five that have the hydrogen atoms recorded but the dSNAP program uses the heavy atoms to calculate the differences and so clusters them independently of the location of the hydrogen atoms. BARAAD [153] has been seen in several clusters and this is its third appearance.

Two structures that contain only buckling (EADBAR10 [146] and QEDFEK [147]), are more similar (indicated by the level of the tie bar connecting the two in the dendrogram) as are the two different hits from the one structure (DETVAA [154]) that both have a slight propeller twist present.

8.3.1.6 Analysis of Cyan Striped, Blue Striped and Purple Striped Clusters

There are three hits that are on their own at the right hand side of the dendrogram and these have been separated for very good reasons. The cyan striped structure contains some single hydrogen bonds between the two defined faces (the defined areas of the two starting molecules in the set up of the analysis) of the adenine and thymine(KECYUN [143]). This structure also appears in the green cluster, but makes a second appearance due to the presence of a single hydrogen bond between the amine group of the purine base and one of the oxygen atoms of the pyrimidine base (Figure 8.16). This hydrogen bond is of moderate strength and has an N...O distance of 2.932 Å. This cluster is on its own as it is the only structure that contains this single hydrogen bond between the two bases. This bond acts as a link between separate base pair motifs in a chain in the structure.

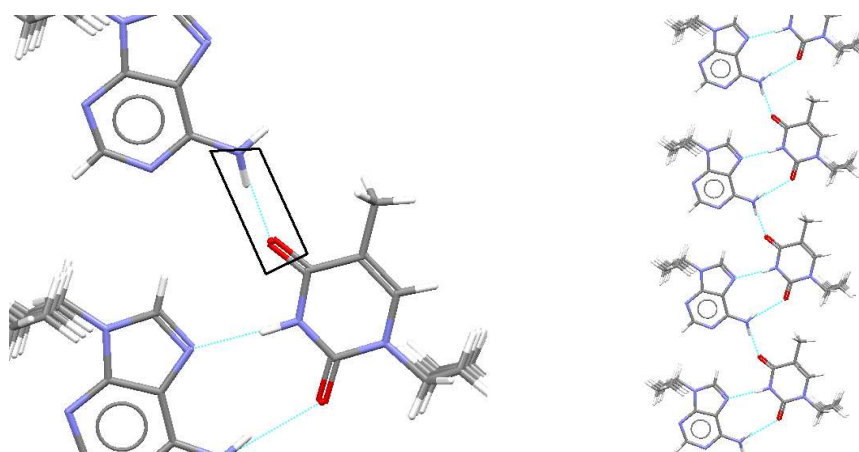


Figure 8.16. The single hydrogen bond in the cyan striped (KECYUN [143]) that extends base paired units into a chain.

The next cluster that contains a single hit is the blue striped cluster (FUJDET [157]) that results from offset base stacking interactions between the two defined faces of the starting molecules (Figure 8,17). These base stacking interactions are of length of approximately 3.2 Å.

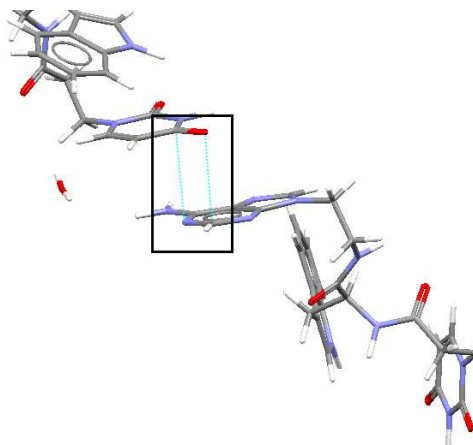


Figure 8.17. Base stacking interactions between the defined areas of the starting molecules in the blue striped cluster (FUJDET [157]).

The KECYUN [143] structure appears in the green cluster as it also has a reverse Hoogsteen base pairing motif. The only other connection between the two defined regions is the edge to edge stacking interactions that are present between parallel layers in this hit. The interactions here are between one of the oxygen atoms of the thymine and the carbon in between the two nitrogen atoms of the six membered ring of the adenine (3.177 Å). The other interaction is between the nitrogen of the adenine next to the amine group and the carbon directly bonded to the previously mentioned oxygen of the thymine (3.242 Å). This is the not only structure that was clustered according to stacking interactions between the two defined faces.

The final cluster that contains only one hit is the purple striped cluster (MIUDAP10 [151]). This structure is the other structure that contains base stacking interactions between the two defined faces (Figure 8.18). The difference between this cluster and the previous is the area involved in the interactions. The first base stacking interaction (1 in Figure 8.18) in this structure is between the one of the carbon atoms of the thymine bonded to an oxygen atom and the nitrogen in between the two amine groups of the adenine molecule (3.233 Å). The second interaction (2) is between the other carbon atom that is bonded to an oxygen atom of the thymine and one of the amine groups of the adenine group (3.509 Å).

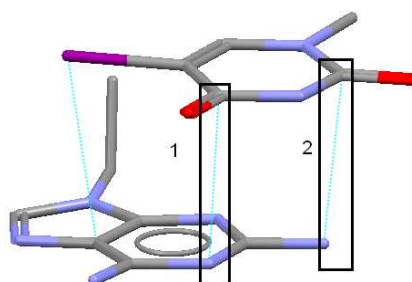


Figure 8.18. Base stacking interactions between the defined areas of the starting molecules in the purple striped cluster (MIUDAP10 [151]).

dSNAP has been shown to be very effective at clustering adenine and thymine fragments according to their base pairing classification. It has not shown the ability to distinguish different levels of propeller twisting and buckling from one another; this affect may be too subtle in the presence of the dominant hydrogen bonding motifs.

8.4 dSNAP Study of Cytosine and Guanine Base Pairing

8.4.1 Construction of the Search

The cytosine and guanine search was conducted in a similar manner to that of the adenine thymine survey with the hydrogen bonding groups of both molecules defined (Figure 8.19) in the Conquest program [172]. This resulted in nine hits.

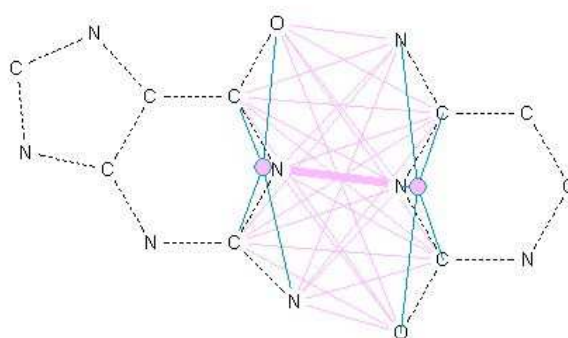


Figure 8.19. Defined contacts used for search in the CSD between cytosine and guanine molecules.

The solubility characteristics of both cytosine and guanine may be the reason for the restricted number of hits found in the CSD. Cytosine is very soluble in methanol and water and guanine is nearly insoluble in these two solvents. The fact that there is no reported molecular complex of cytosine and guanine may indicate how difficult it can be to obtain these structures in a complex.

The similarity level calculated by the dSNAP program led to five clusters with three clusters only containing one hit (Figure 8.20). The MMDS plot clearly shows that the yellow hit should be included in the red cluster (Figure 8.21) which is also indicated by the close proximity of the tie bar connecting the yellow and red clusters and the cut level. The cut level was therefore decreased to create a single cluster generated from a merging of the red and yellow clusters.

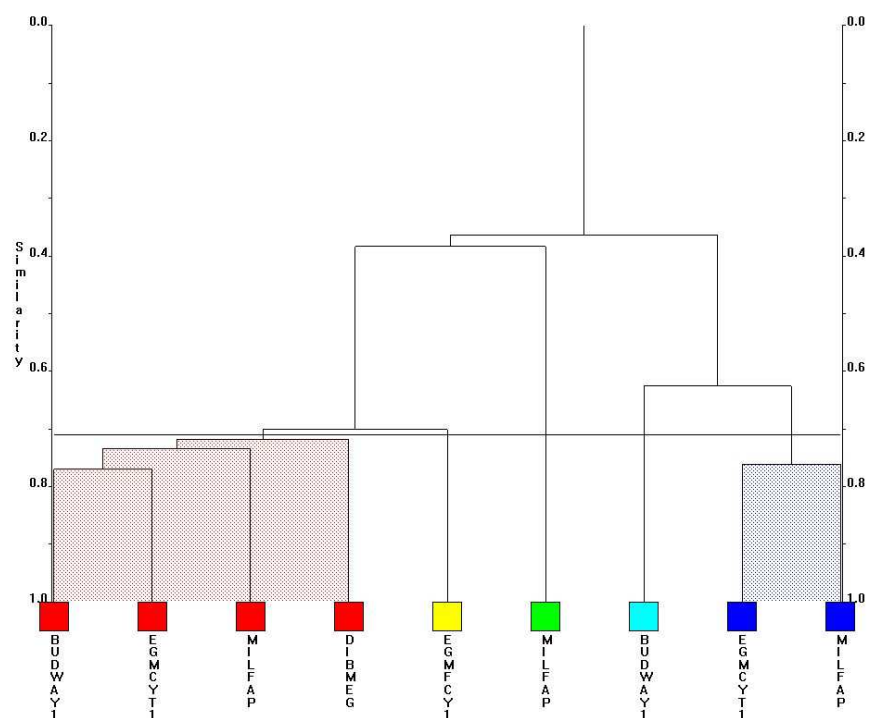


Figure 8.20. Dendrogram obtained from the dSNAP program when with the automatic cut level for cytosine and guanine fragments.

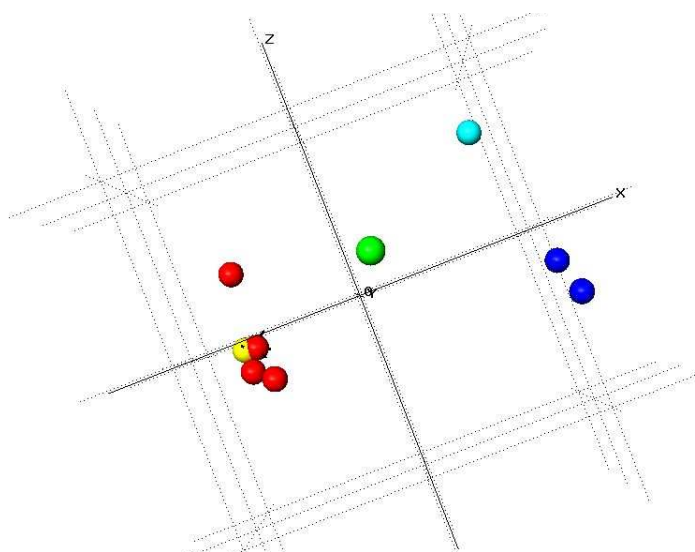


Figure 8.21. MMDS plot of results obtained from the dSNAP program using the calculated cut level for the cytosine guanine hydrogen bonded fragments.

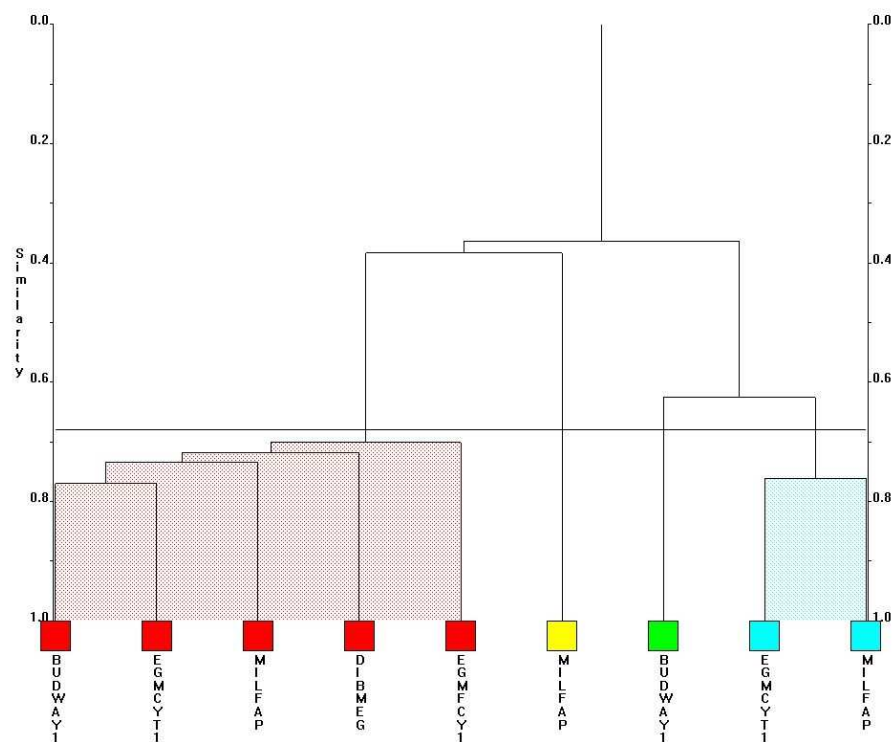


Figure 8.22. Dendrogram with adjusted cut level to move one hit into the red cluster.

After the adjustment of the cut level, four clusters were obtained with two clusters only containing one hit.

8.4.2 Analysis of the Red Cluster

The red cluster contains base pair motifs of a standard Watson-Crick bonding motif between cytosine and guanine. This is made up of three hydrogen bonds between the cytosine and guanine molecules (Figure 8.23).

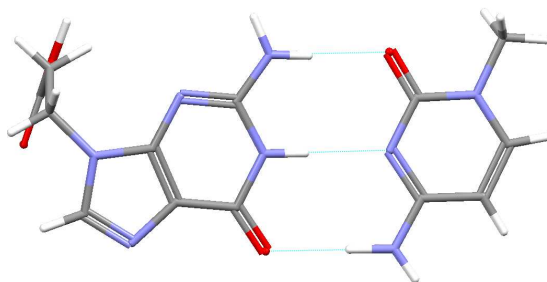


Figure 8.23. Primary hydrogen bonding motif exhibited by all members of the red cluster illustrated by BUDWAY10 [160].

Table 8.7. Bond distances of hydrogen bonds for the base pair between the cytosine and guanine in the red cluster. Going down column one represents going from left to right in the cluster

Bond	H-N---O (Å)	N---H-N (Å)	O---H-N (Å)
1. BUDWAY10 [160]	2.800	2.938	2.982
2. EGMICYT10 [161]	2.806	2.924	2.943
3. MILFAP [162]	2.942	2.947	2.867
4. DIBMEG [162]	2.788	2.915	2.921
5. EGMFCY [131]	2.822	2.942	2.964

There is a large degree of variation in the hydrogen bond distances as the shortest distance is not always in the same position in the fragment. The main driving force behind the formation of this cluster is therefore the hydrogen bonding motif as all of the structures show Watson-Crick pairing which is the naturally occurring bonding motif between these two. This is also the most common hydrogen bonding motif found in the CSD between guanine and cytosine units.

8.4.3 Analysis of Remaining Clusters

The fragment in the yellow cluster (MILFAP [162]) is also found in the red cluster. The reason for this being clustered on its own is therefore a second single hydrogen bond that is not seen in any other cluster. This single hydrogen bond is between the amine group of the cytosine molecule and one of the nitrogen atoms in the five membered ring of the guanine molecule (Figure 8.24).

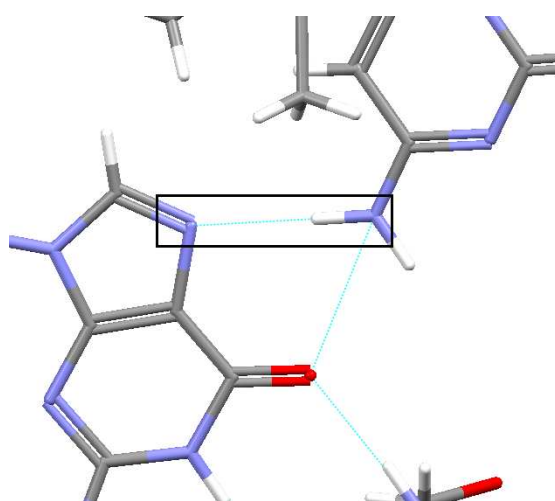


Figure 8.24. Single hydrogen bond giving rise to the yellow cluster MILFAP [162].

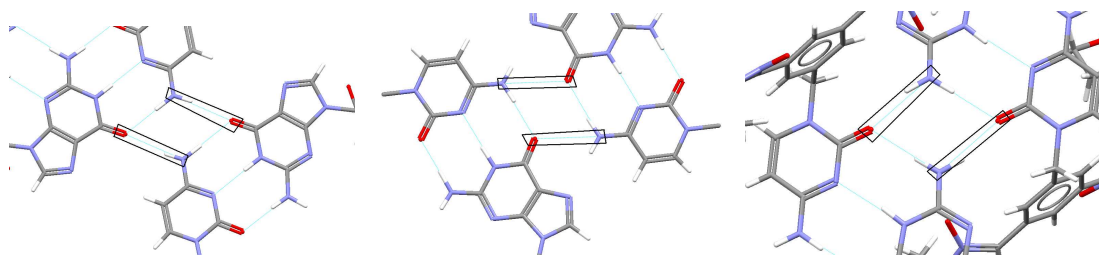


Figure 8.25. Similar hydrogen bonding motif in the hit from the green cluster; BUDWAY10 [160], left and the two members of the cyan cluster, middle EGMCYT10 [161] and right MILFAP [162].

The last remaining cluster that contains more than one hit is the cyan cluster. These fragments also occur within the red cluster. The reason the fragments cluster into a second grouping is because of single hydrogen bonds that connect Watson Crick units of cytosine and guanine together (Figure 8.25). In each case the link between cytosine and guanine is between the amine group of the cytosine and the oxygen atom of the guanine.

The green cluster has single hydrogen bonds that are very similar to those found in the cyan structure with a slightly different orientation (Figure 8.25, left). This hit may therefore be better included in the cyan cluster. This means that the threshold line would be best adjusted to include the green cluster in the cyan cluster resulting in three clusters (Figure 8.26). After the adjustment of the threshold line the most sensible dendrogram has been gained that is made up of three clusters with one cluster containing only one hit. This clustering does make sense but this may require more hits to test the interpretation of the clustering further.

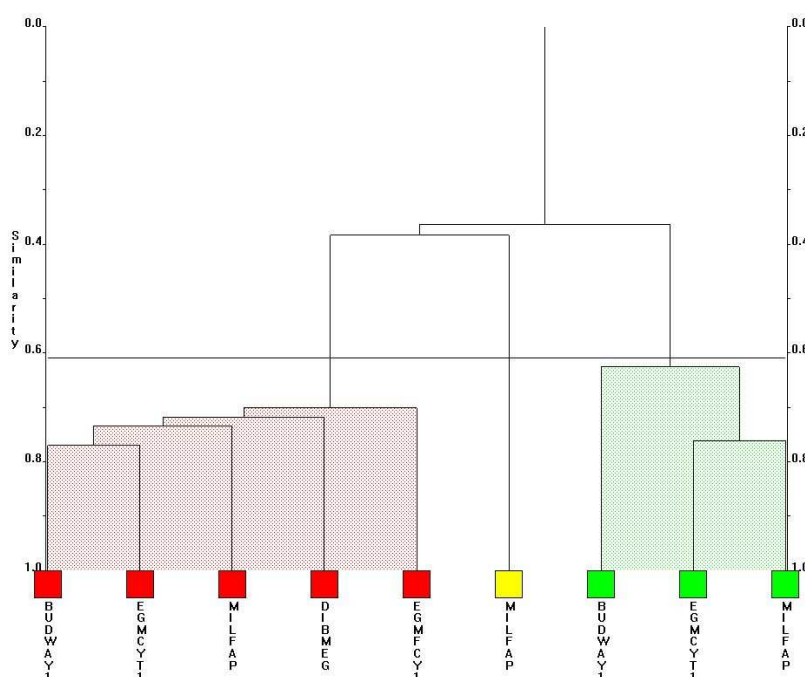


Figure 8.26. Dendrogram after second movement of the similarity cut-off to incorporate the cyan and green clusters into one green cluster.

8.5 CSD Survey of Cytosine Base Pairing

A search of the CSD was undertaken to identify the common hydrogen bonding motifs previously reported for the cytosine molecule. The search was restricted to include only cytosine and protonated cytosine molecules and no cytosine derivatives. All of the bonds within the cytosine molecule were defined as any bond type (Figure 8.27) to ensure that those showing hydrogen transfer were also included in the survey. Any hits corresponding to cytosine derivatives were omitted from the search results manually. This left 53 hits. The results from the CSD are compared with the molecular complexes reported in this thesis.

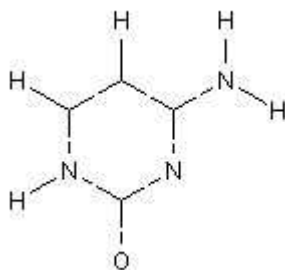


Figure 8.27. Search fragment defined for cytosine molecular complexes in the CSD.

8.5.1 Analysis of the Hydrogen Bonding Patterns Found in the CSD

The first criteria for classifying the cytosine based molecular complexes contained within the CSD is whether or not hydrogen transfer to a cytosine has occurred creating a charged cytosinium molecule. The transfer of a hydrogen atom can have a profound affect on the hydrogen bonding that materialises in the structure. Of the 53 hits found in the CSD, 31 of these had 100% charged cytosinium molecules, 15 showed both neutral and charged cytosine molecules and 7 contained only neutral cytosine molecules in the molecular complexes. 32 cytosine molecular complexes are reported in this thesis; of these 13 of these had 100% charged cytosine molecules, 11 showed both neutral and charged cytosine molecules and 8 contained only neutral cytosine molecules in the molecular complexes. The relative occurrence of each of these is thus consistent between those contained within the CSD and those reported in this thesis.

8.5.2.1 Protonated Cytosine Molecules Only

The presence of only protonated cytosine molecules restricts the possible hydrogen bonding patterns available. From the CSD, there are 31 hits that have total hydrogen transfer. There is no occurrence of the pseudo-Watson-Crick motif or of the type A hydrogen bonding motif. The type B hydrogen bonding motif is present in four of the

structures. However, the most prominent hydrogen bonding motif is the hydrogen-bonded heterodimer ring with 15 of the structures exhibiting it. The remaining structures exhibit no primary bonding motif (13 structures).

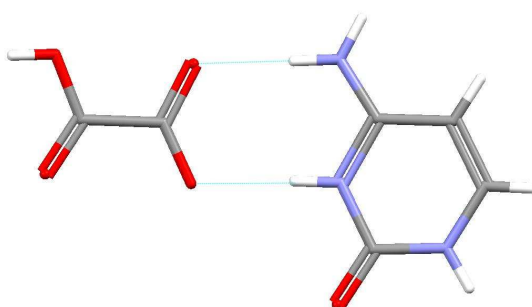


Figure 8.28. Heterodimer hydrogen bonded ring formed once a hydrogen atom is transferred to the cytosine molecule illustrated by CASCIJ [173].

15 of the 33 CSD hits show a hydrogen bonded heterodimer (Figure 8.28). All 13 of the molecular complexes that exhibit total hydrogen transfer reported in this thesis show these hydrogen-bonded heterodimers. This hydrogen bonding motif is therefore the most commonly occurring hydrogen bonding motif. The type B hydrogen-bonding motif is able to form concurrently with that of the hydrogen bonded heterodimer or with the type A and pseudo-Watson-Crick. However, 13 molecular complexes in the CSD do not show any of these hydrogen-bonding motifs.

In conclusion, the primary hydrogen bonding motif for a molecular complex containing only protonated cytosinium molecules is the formation of hydrogen-bonded heterodimers with the possibility of a secondary hydrogen bonding motif in the form of a type B bonding motif also occurring.

Table 8.8. The hydrogen bonding motifs found in charged cytosine molecular complexes.

	Total Number with Total Hydrogen Transfer	Number containing pseudo-Watson-Crick	Number containing Type A	Number containing type B	Number containing hydrogen-bonded ring
Number in CSD	31	0	0	4	15
Number in Thesis	13	0	0	3	13

8.5.2.2. Mixed Charge Cytosine Molecular Complexes

In this group, there is hydrogen transfer to the unprotonated ring nitrogen of a cytosine molecule, but this protonated molecule coexists alongside neutral cytosine molecules. The pseudo-Watson-Crick hydrogen-bonding motif is therefore possible (Figure 8.29).

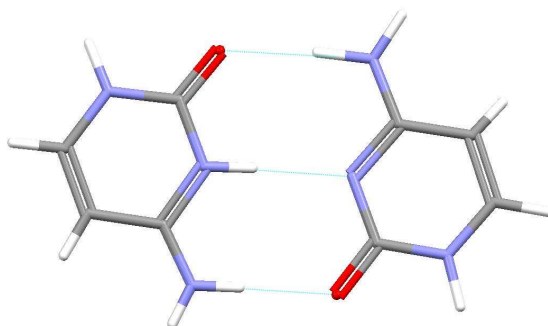


Figure 8.29. Pseudo-Watson-Crick base pair which can only be formed in the presence of both neutral and protonated cytosine molecules shown by CYTZNC [174].

All 15 structures found in the CSD which contain both protonated and neutral cytosine molecules display the pseudo-Watson-Crick hydrogen-bonding motif. Out of these, this is the only motif found in 10 of these complexes and 5 contain the type B hydrogen bonding motif in addition, this pattern is also found when considering the structures reported in this thesis; all 11 structures show this pseudo-Watson Crick hydrogen bonding motif. Three complexes also show the type B hydrogen bonding motif but the majority form hydrogen bonded chains of pseudo Watson Crick dimers connected through pairs of single hydrogen bonds (Figure 8.30). This is a common occurrence for structures that contain a pseudo-Watson-Crick motif. The hydrogen bonded ring motif appears in one structure (4-hydroxybenzoic acid and cytosine), however, this is uncommon and the hydrogen bonded ring motif terminates the extended pseudo-Watson-Crick chain from forming.

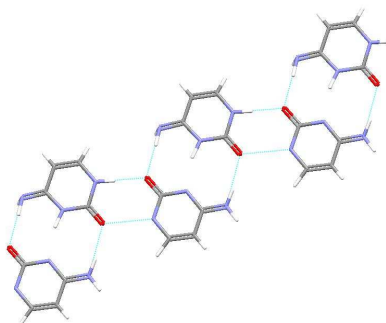


Figure 8.30. Pseudo-Watson-Crick base pairs linked via type B bonding motifs represented by CYTRES10 [169].

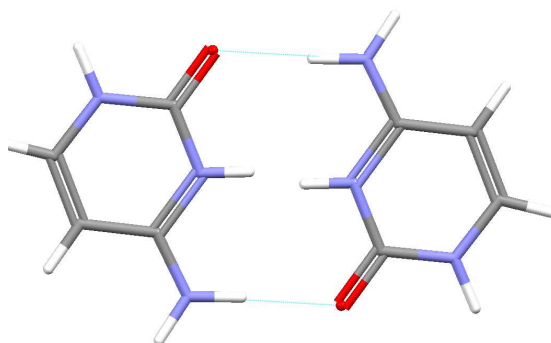


Figure 8.31. Two hydrogen atoms present either side of an inversion centre with each having an occupancy factor of 0.5 represented by ODICOU [163].

One compound of interest (ODICOU) is cytosinium cytosine 7,7,8,8-tetracyanoquinodimethane[163]. In the literature the hydrogen bond within the central hydrogen bond is found to be disordered over two positions. This is also found in TAZXAU [164], TAZXEY [164], and LIWSIU [165].

Table 8.9. The number of structures in the CSD and this thesis that show mixed cytosine and cytosinium molecular complexes and the hydrogen bonding motifs found.

	Total Number with Partial Hydrogen Transfer	Number containing pseudo-Watson-Crick	Number containing Type A	Number containing type B	Number containing hydrogen-bonded ring
Number in the CSD	15	15	0	5	0
Number in thesis	11	11	0	3	1

The CSD results shows the dominance of the pseudo-Watson-Crick bonding motif, which is present in all of the structures. The only other primary bonding motif seen is the type B bonding motif and it does not use the same face of the cytosine as the pseudo-Watson-Crick and so this forms as a secondary bonding motif. The type B bonding motif results in the formation of extended cytosine chains.

8.5.2.3 Neutral Cytosine Molecules Only

In the CSD, there are seven molecular complexes involving just pure cytosine (not including 5-fluorocytosine) that do not contain any hydrogen transfer. This means that the structures are not capable of forming the three hydrogen bonds required to make the pseudo-Watson-Crick and as a result this bonding motif is not seen in any of the structures. These structures are capable of forming the type A, type B, type C and heterodimer motifs.

The structures in this group do not show a predominance of any of the hydrogen bonding motifs reported in this thesis. Only one molecular complex shows the type B hydrogen-bonding motif. Two structures do not exhibit any primary hydrogen bonding motifs, but both of these are metal complexes, and the first has a protonated amine group (BADFOC [166] and CYTSCA [167]). The remaining four structures all form a different hydrogen bonding motif which is a hybrid of the type A and B bonding motifs, type C (Figure 8.32).

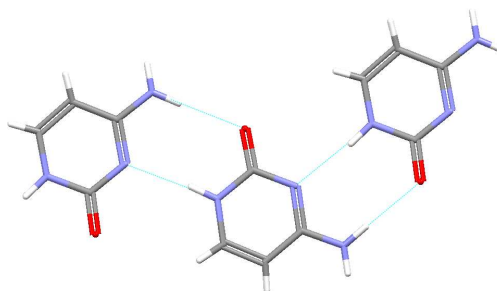


Figure 8.32. Type C hydrogen bonding motif found in several of the CSD examples illustrated by CYTOSM [170].

Table 8.10. Number of structures from CSD and thesis that contain no hydrogen transfer and the hydrogen bonding motifs found

	Total number with no H transfer	Number containing pseudo-Watson-Crick	Number containing type A	Number containing type B	Number containing hydrogen-bonded ring
Number in CSD	7	0	0	1	0
Number in thesis	8	0	7	8	1

The structures reported in this thesis show a dominance of the type A and B hydrogen bonding motifs with only one structure showing the hydrogen bonded ring motif (5-fluorocytosine and 4-hydroxybenzoic acid). However, the CSD searches show no such dominance but show a different hydrogen bonding motif, a hybrid of the type A and B bonding motifs, type C, which appears in four out of the seven structures.

8.5.3. Summary

The most common state for the cytosine based on the molecular complexes found both within the CSD and this thesis is the fully protonated form with 45% of the structures in this thesis and 58% of the structures in the CSD (Table 8.11) showing the fully protonated cytosine molecule (Figure 8.33) with no neutral cytosine molecule. The next most common is the mixed neutral and protonated cytosine species. Neutral complexes are the less frequently reported.

Table 8.11. Comparison of the prevalence of hydrogen transfer in the CSD and thesis structures.

	CSD – number of structures	CSD % of total number of results	Thesis – number of structures	Thesis structures - % of total
Protonated cytosine	31	58	13	41
Mixed protonated and neutral cytosine	15	28	11	34
Neutral cytosine	7	14	8	25

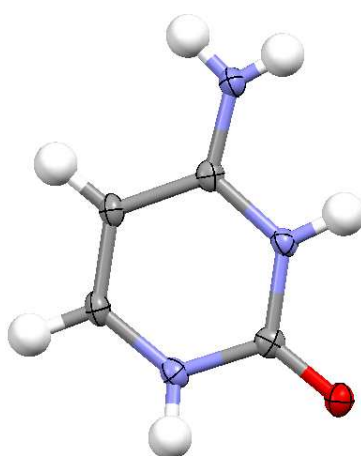


Figure 8.33. Most common state for a cytosine molecule to exist in the CSD represented by ADALAS [171].

The hydrogen bonding motifs found in these molecular complexes are thus heavily dependent on the level of hydrogen transfer. More than one bonding motif was found to occur within the same complex, but one dominates and the other acts as a connecting bond. The most common connecting unit is the type B hydrogen-bonding motif. Type B is the only hydrogen-bonding motif that can occur in the presence of other motifs.

Table 8.12. Comparison of the prevalence of the hydrogen bonding motif in the CSD and thesis structures.

	CSD structures – number showing bonding motif	% of structures showing bonding motif in CSD	Thesis structures – number showing bonding motif	% of structures showing bonding motif in thesis
pseudo- Watson-Crick	15	29	11	34
Type A	0	0	7	21
Type B	10	20	14	43
Hydrogen- bonded heterodimer	15	29	15	47
Alternative hydrogen bonding motif	15	29	0	0

In the CSD structures, the relative occurrence of the five different hydrogen-bonding motifs is very similar. The structures reported in this thesis show only the three different types and the most common hydrogen-bonding motif is the hydrogen-bonded heterodimer. The pseudo-Watson-Crick and type A and B bonding motifs occur in approximately equal amounts.

9. Conclusions

A diverse range of molecular complexes of nucleobases, primarily cytosine and 5-fluorocytosine have been generated. Some common hydrogen bonding motifs were apparent and these are primarily influenced by the level of proton transfer in the molecular complexes. The effect of fluorinating both the nucleobase and the co-molecule has been examined and found not to disrupt the base pairing motifs, rather affecting the more subtle, weaker intermolecular interactions. In some cases, these intermolecular interactions only introduced a small change in the global packing while in other cases, the crystal packing completely changed on fluorination of one or both components.

9.1. Molecular Complexes of Cytosine and 5-Fluorocytosine

The nucleobase cytosine and its fluorinated equivalent 5-fluorocytosine both contain an unprotonated N atom in the heterocyclic ring. This nitrogen is capable of accepting a hydrogen atom transferred from an acid co-molecule during molecular complex formation. Uracil and 5-fluorouracil do not contain an unprotonated heteroatom and so their molecular complexes are not influenced by proton transfer effects.

9.1.1 pK_a Matching

The ΔpK_a values have, in general, provided a good rationale for explanation of the hydrogen transfer effects resulting from the co-crystallisation attempts between the nucleobases and their co-crystallisations with acids. The ratio of crystallisation and the presence of water have also been seen to have an effect on the hydrogen transfer. In general, cytosine molecular complexes showing fully charged species resulted in 1:1 molecular complexes. Those that crystallised in a 2:1 cytosine to co-molecule ratio showed only partial hydrogen transfer.

The molecular complexes reported in this thesis can be broadly split into three categories when referring to pK_a matching: 1:1 ratio, not 1:1 ratio, and hydrates or solvates.

9.1.1.1 1:1 Ratio

Within the class of 1:1 molecular complexes, all three proton transfer classes are obtained: 100%, partial and no proton transfer, and the molecular complexes obtained follow the pK_a rules suggested by Childs *et al* (Section 1.6) [49]. The three molecular complexes that

show no hydrogen transfer (5-fluorocytosine with benzoic acid, 2-fluorobenzoic acid and 4-hydroxybenzoic acid) all have a ΔpK_a value of less than 0 (Table 9.1).

The remaining eight structures all have ΔpK_a values falling into the continuum region where it is difficult to predict whether or not proton transfer will occur. The majority of these show total proton transfer to the cytosine/5-fluorocytosine molecule. Of the two that show only partial proton transfer, the 5-fluorocytosine and 3,5-dinitrobenzoic acid molecular complex shows disordered hydrogen atoms.

Table 9.1 ΔpK_a values for the cytosine or 5-fluorocytosine molecular complexes that resulted in a 1:1 molecular complex ratio

Molecular Complex	ΔpK_a	Proton Transfer
Cytosinium 3-Fluorobenzoate	0.74	Total
Cytosinium Salicylate	1.62	Total
Cytosinium 5-Fluorosalicylate	1.92	Total
5-Fluorocytosium 2,6-Dihydroxybenzoate Form I	1.96	Total
5-Fluorocytosinium 2,6-Dihydroxybenzoate Form II	1.96	Total
Cytosine and Benzoic Acid	0.396	Partial
5-Fluorocytosine and 3,5-Dinitrobenzoic Acid	0.76	Partial
5-Fluorocytosine and Benzoic Acid	-0.944	None
5-Fluorocytosine and 2-Fluorobenzoic Acid	-0.01	None
5-Fluorocytosine and 4-Hydroxybenzoic Acid	-1.31	None
Salicylic acid and 5-Fluorocytosine	0.28	Total

9.1.1.2 Anhydrous Non 1:1 Ratio Molecular Complexes

The next group of molecular complexes are those that form in a non 1:1 stoichiometric ratio and do not contain any solvent molecules. All three of the molecular complexes in this class crystallise in a 2:1 ratio of cytosine to co-molecule and show only partial proton transfer, except for the 3-hydroxybenzoic acid and 5-fluorocytosine structure that shows no hydrogen transfer. As shown in Section 1.6 the crystallisation ratio of the product can have an effect on how applicable the ΔpK_a values are on the predictability of hydrogen transfer. In the case of two of these molecular complexes, the ΔpK_a values fall within the region between 0 and 3 making it difficult to predict the level of proton transfer. The presence of both charged and neutral species in these complexes is consistent with this. The remaining structure has a ΔpK_a value of -0.82 which would mean hydrogen transfer is unfavourable and there is no hydrogen transfer exhibited.

Table 9.2 ΔpK_a values for the cytosine or 5-fluorocytosine molecular complexes not in 1:1 ratio and containing no solvent

Molecular Complex	ΔpK_a	Proton Transfer
Cytosine and 2-nitrobenzoic acid	2.44	Partial
Cytosine and 3-bromobenzoic acid	0.79	Partial
3-hydroxybenzoic acid and 5-Fluorocytosine	-0.82	None

9.1.1.3 Molecular Complexes Containing Solvent Molecules

The final class of structures are those complexes that contain solvent. The most common solvent to appear in the structures was the water molecule. All three levels of protonation state are found in complexes in this class.

Table 9.3 ΔpK_a values for the cytosine or 5-fluorocytosine molecular complexes not in 1:1 ratio but containing solvent

Molecular Complex	ΔpK_a	Proton Transfer
Cytosinium and 3-Hydroxybenzoate Hemihydrate	0.52	Total
Cytosinium 2,4-Dihydroxybenzoate Hydrate Form I and II	1.49	Total
5-Fluorocytosinium 3,5-Dihydroxybenzoate Hydrate	-0.78	Total
Cytosinium 3,5-Dihydroxybenzoate Hydrate	0.56	Total
Cytosinium Nicotinate Monohydrate	2.6	Total
Cytosinium 2,4,6-trimethylbenzoate monohydrate	1.17	Total
Cytosine and 2-Fluorobenzoic acid Hydrate	1.33	Partial
Cytosine and 4-Fluorobenzoic acid Hydrate	0.45	Partial
5-Fluorocytosine and 5-Fluorosalicic Acid Hydrate	0.58	Partial
Cytosine and 5-Fluorosalicic Acid Hydrate	1.92	Partial
Cytosine and 4-Hydroxybenzoic Acid Hydrate	0.03	Partial
Cytosine and 2,6-Dihydroxybenzoic Acid Hydrate	3.30	Partial
Cytosine and Isonicotinic acid Hydrate	2.83	Partial
5-fluorocytosine and 3-fluorobenzoic acid	-0.600	None
5-fluorocytosine and 4-fluorobenzoic acid	-0.890	None

In general, the ΔpK_a values for the molecular complexes in this class fall in the region between 0 and 3 and thus a varying degree of protonation is obtained. The main molecular complexes to deviate from the rules are 5-fluorocytosinium 3,5-dihydroxybenzoate hydrate and cytosine and 2,6-dihydroxybenzoic acid hydrate. The 5-fluorocytosinium 3,5-dihydroxybenzoate hydrate molecular complex has a ΔpK_a value of -0.78 meaning that hydrogen transfer is unfavourable and unexpected, however, total hydrogen transfer is

exhibited. Contrastingly, the ΔpK_a value for the cytosine 2,6-dihydroxybenzoate molecular complex is 3.30 suggesting that full proton transfer should be expected. However, in this case, only partial proton transfer is obtained. The introduction of a water molecule or the non 1:1 ratio of the products may result in the diminished effectiveness of the pK_a rule.

9.1.2 Base-Pair Motifs

Hydrogen transfer has a significant effect on the primary bonding motif adopted in the molecular complexes reported in this thesis. There are three possible categories for the molecular complexes involving cytosine and 5-fluorocytosine, the structures capable of hydrogen transfer: total hydrogen transfer (all cytosine or 5-fluorocytosine molecules are protonated), partial hydrogen transfer (a mix of charged and neutral cytosine or 5-fluorocytosine molecules), and no hydrogen transfer (only neutral cytosine molecules present in the structure).

9.2 Hydrogen Transfer

9.2.1 Total Hydrogen Transfer

Table 9.4 Total number of cytosine or 5-fluorocytosine molecular complexes with protonated, mixed protonated states and neutral only cytosine molecules.

	Thesis – number of structures	Thesis structures - % of total
Protonated cytosine	13	40
Mixed protonated and neutral cytosine	11	35
Neutral cytosine	8	25

Thirty two molecular complexes including either cytosine or 5-fluorocytosine have been reported in this thesis. From these structures just under half (13) exhibited total hydrogen transfer. When a structure exhibits total hydrogen transfer it limits the bonding patterns that can be adopted by the structure as the pseudo-Watson-Crick bonding motif is not possible and neither is the type A bonding motif. In all of the structures that contain total hydrogen transfer the hydrogen-bonded ring motif was formed (Figure 9.1 (left)).

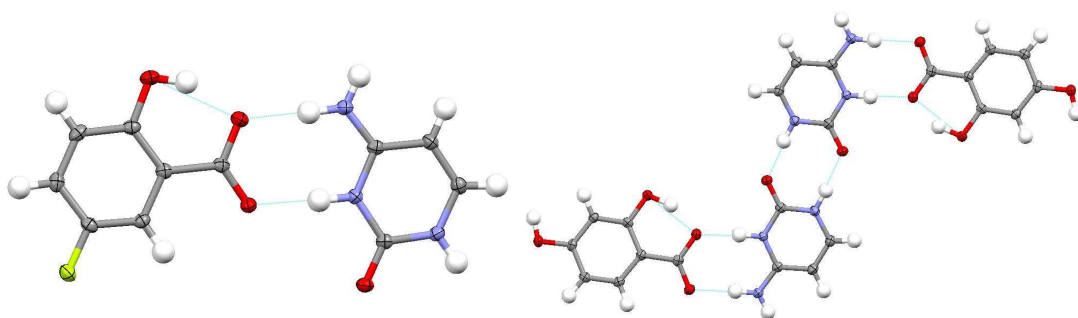


Figure 9.1. Hydrogen bonded ring motif exhibited by the molecular complex of cytosinium 5-fluorosalicylate (left) and the supramolecular unit exhibited by the cytosine and 2,4-dihydroxybenzoic acid hydrate molecular complex (right).

This bonding motif involves the two areas of the co-molecules that have either lost or gained the hydrogen and as such this bonding motif may be favoured here to balance the charge occurring as a result of the transfer of the hydrogen from the co-molecule to the cytosine or 5-fluorocytosine molecule.

The type B bonding motif is present in three of the thirteen structures. This bonding motif helps to create a small building block unit (Form I and II of cytosine with 2,4-dihydroxybenzoic acid hydrate) as seen in Figure 9.1 (right). However, the more common bonding pattern was for there to be only single hydrogen bonds from the original hydrogen-bonded ring pair to neighbouring hydrogen-bonded ring pairs as seen in the cytosine and 5-fluorosalicyclic acid structure and shown in Figure 9.2. This total hydrogen transfer gives a predisposition to the structure to form the hydrogen-bonded ring motif as it was adopted by all of the structures shown in this thesis that contain total hydrogen transfer.

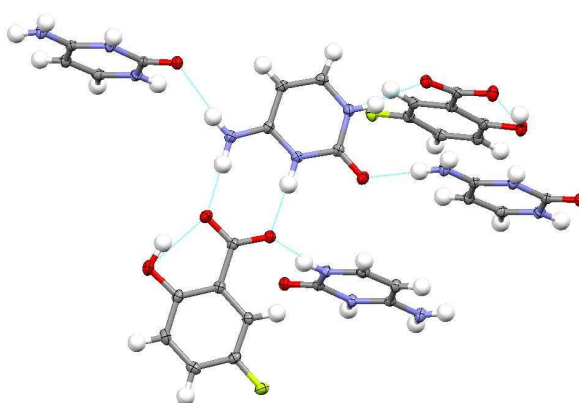


Figure 9.2. Common bonding pattern with only single hydrogen bonds linking hydrogen bonded ring units together exhibited by the molecular complex of cytosinium 5-fluorosalicylate.

9.2.2 Partial Hydrogen Transfer

Eleven molecular complexes containing only partial hydrogen transfer with respect to the cytosine/5-fluorocytosine molecule were identified. This partial hydrogen transfer means that there are both charged and neutral cytosine or 5-fluorocytosine molecules in the structure. The co-molecule in these molecular complex could either exist solely in its charged state (in 2:1 molecular complexes) or with both charged and neutral co-molecules within the molecular complex (1:1 or other ratios).

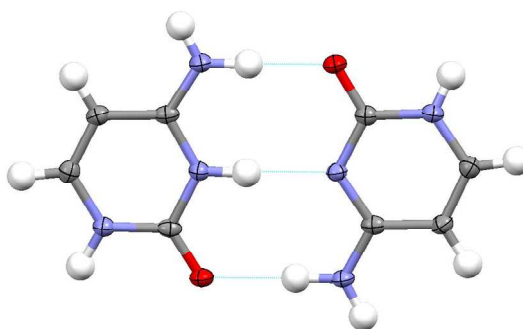


Figure 9.3. pseudo-Watson-Crick base pair motif exhibited by the molecular complex of cytosine and 4-fluorobenzoic acid hydrate.

All eleven structures that possessed partial hydrogen transfer exhibited the pseudo-Watson-Crick base pairing motif shown in Figure 9.3. This bonding motif is made up of three hydrogen bonds between two cytosine molecules (one charged and one neutral). This bonding motif is only possible when both charged and neutral cytosine or 5-fluorocytosine molecules are present. Once again the hydrogen transfer gives the structure a predisposition to forming the pseudo-Watson-Crick bonding motif and this may once again be due to the balancing of the charge in the cytosine molecule that has become protonated.

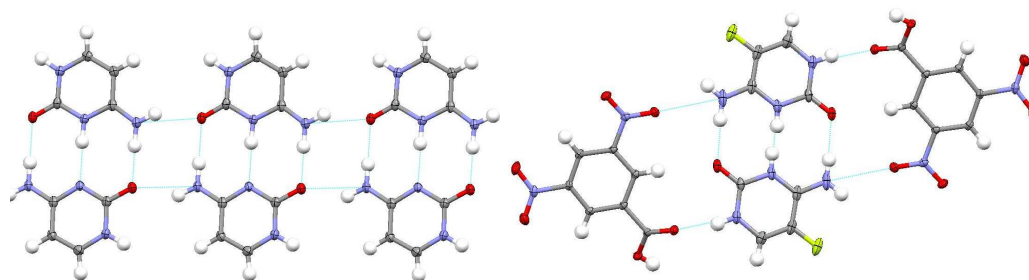


Figure 9.4. Two possible bonding patterns – extended cytosine homo-base-paired chain as exhibited by the molecular complex of cytosine and 2-fluorobenzoic acid hydrate (left) and the formation of a supramolecular unit as seen in the molecular complex of 5-fluorocytosine and 3,5-dinitrobenzoic acid (right).

In the majority of the molecular complexes, the pseudo-Watson-Crick units form into extended chains via single hydrogen bonds between adjacent pairs as shown in Figure 9.4 (left). In some of the structures (e.g. 5-fluorocytosine and 3,5-dinitrobenzoic acid) the

chain is terminated by the co-molecule to give a basic building block unit in the structure rather than the extended chain (Figure 9.4 (right)). These basic building blocks can be linked to one another via single hydrogen bonds. Of the structures determined, 11 exhibit the pseudo-Watson-Crick bonding motif. Of these 11 structures seven exhibit the pseudo-Watson-Crick extended chain as seen in Figure 9.4 left. The remaining four structures all have their chains terminated by the co-molecule present in the complex.

9.2.3 No Hydrogen Transfer

The remaining eight structures all exhibit no hydrogen transfer. This category simply refers to a structure made up of only neutral cytosine molecules and neutral co-molecules. The molecules in this category once again are not capable of forming the pseudo-Watson-Crick bonding motif but are capable of forming the hydrogen-bonded ring, type A and type B bonding motifs. All of the structures show the type B bonding motif with this acting as a link between either hydrogen-bonded ring motifs or type A bonding motif. Type B does not require the face of the cytosine molecule that is involved in all the other bonding motifs and as a result there is no restriction to the type B bonding motif forming here.

In seven of the eight structures, extended homo base-paired chains are formed between cytosine/5-fluorocytosine molecules and these are comprised of alternating type A and type B bonding motifs (Figure 9.5, left). The chains are then linked to one another through the co-molecule.

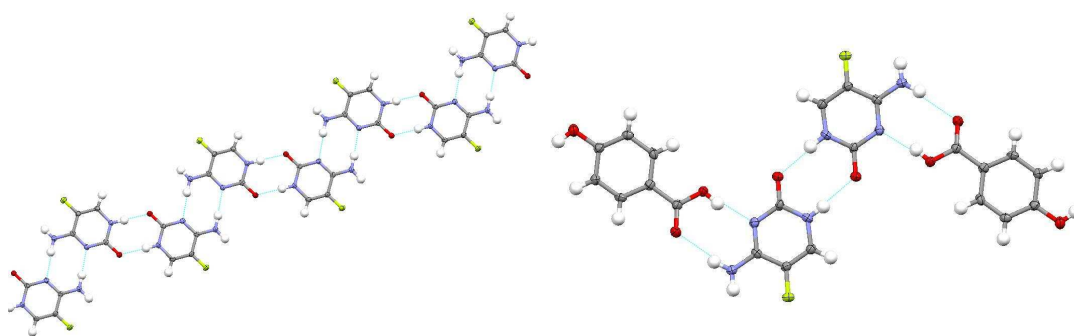


Figure 9.5. The extended homo base-paired chain formed by alternating type A and type B bonding motifs in the molecular complex of 5-fluorocytosine and 3-fluorobenzoic acid solvate (left) and the terminated chain seen in the molecular complex of 5-fluorocytosine and 4-hydroxybenzoic acid.

In the final structure, the molecular complex of 5-fluorocytosine and 4-hydroxybenzoic acid hydrate, this chain is terminated by a combination of the co-molecule and the solvent molecules. This structure only contains the type B bonding motif and does not form the type A bonding motif. The 4-hydroxybenzoic acid molecule terminates the chain in the position that the type A bonding motif would form creating a 4 molecule supramolecular unit (Figure 9.5 (right)).

9.3 Fluorine interactions

The effect of fluorination was investigated in two ways: the effect of fluorination of both the nucleobase itself and the fluorination of the co-molecule was investigated. In some of the structures the fluorine atom plays a pivotal role in the structure whilst in others it merely helps to have subtle effect on the structure that differentiates it from the non-fluorinated equivalent.

9.3.1 Fluorobenzoic acid Molecular Complexes

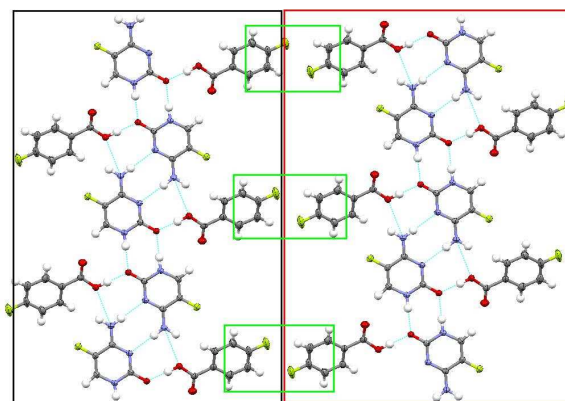


Figure 9.6. Separate chains linked via weak hydrogen bonds in the molecular complex of 5-fluorocytosine and 4-fluorobenzoic acid (green represents the hydrogen bonds that make use of the fluorine atoms and the red and black represent separate chains).

Fluorine hydrogen bonds were found in several of the molecular complexes; the most interesting role of these weaker interactions was found in the molecular complexes of 5-fluorocytosine with 4-fluorobenzoic acid and 3-fluorobenzoic acid. Fluorine hydrogen bonds are formed between pairs of 3/4-fluorobenzoic acid molecules of separate chains creating base-pair type units. These interactions are key to the creation of the interesting architectures in these complexes – the presence of substantial pores.

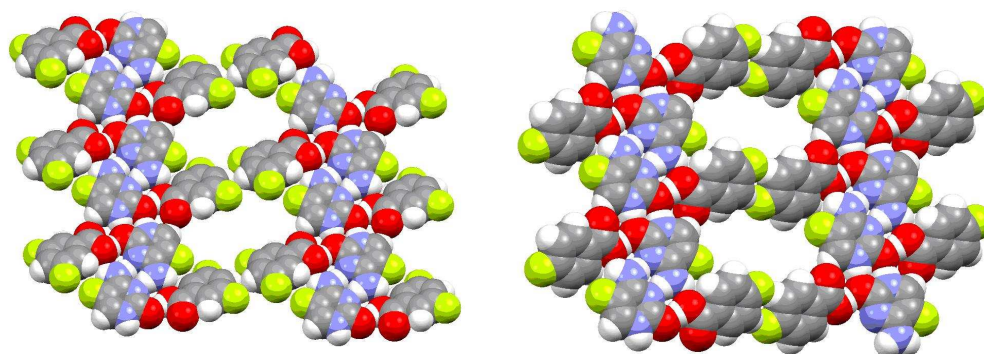


Figure 9.7. Spacefill representation to show the voids present in the molecular complexes of 3-fluorobenzoic acid and 5-fluorocytosine (left) and 4-fluorobenzoic acid and 5-fluorocytosine (right).

In both structures of this type, solvent is contained in the pores. These pores were investigated by a combination of TGA and DSC. In the case of the 5-fluorocytosine : 3-

fluorobenzoic acid molecular complex, these suggested that it might be possible to remove the solvent before the compound decomposes although this requires further investigation. In the case of the 5-fluorocytosine 4-fluorobenzoic acid molecular complex, the outcomes of the thermal analyses were inconclusive.

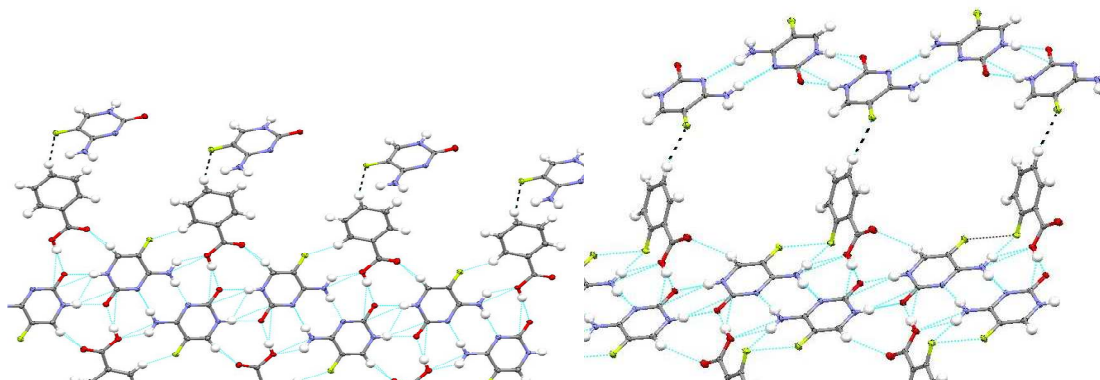


Figure 9.8. Single F \cdots H-C hydrogen bonds connecting chains in the structures of 5-fluorocytosine and benzoic acid (left) and 5-fluorocytosine and 2-fluorobenzoic acid (right)

The same primary homo-base paired chain is evident in all of the complexes synthesized by combining 5-fluorocytosine with benzoic acid, 2-, 3- and 4-fluorobenzoic acid. However, the positioning of the fluorine atom on the benzoic acid affects the possible interactions that can take place. In the benzoic acid molecular complex, the main fluorine interactions present are F \cdots H-C weak hydrogen bonds linking separate chains (Figure 9.8, left). The 2-fluorobenzoic acid molecular complex also contains these via F \cdots H-C hydrogen bonds but invoking a different C-H group from that used in the benzoic acid molecular complex. This is a consequence of F \cdots F interactions between the 5-fluorocytosine and the 2-fluorobenzoic acid molecules within a chain (Figure 9.8, right).

9.3.2 Hydroxybenzoic acid Molecular Complexes

Three related complexes in this family were obtained in the form of cytosinium salicylate, cytosinium 5-fluorosaliclyate and 5-fluorocytosinium salicylate. The differences between these structures are due to the differing positions of the fluorine atoms and hence the fluorine interactions.

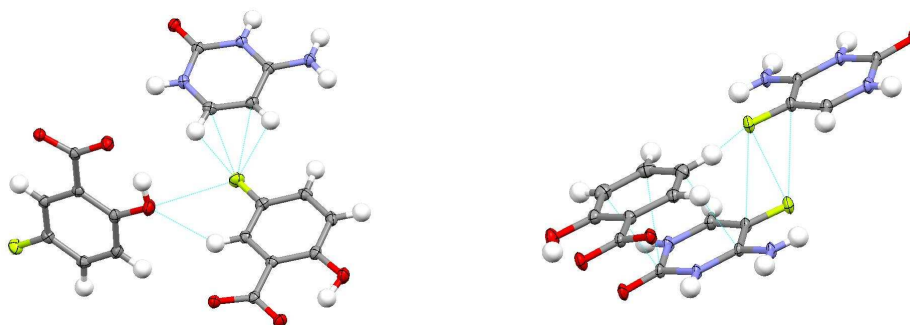


Figure 9.9. Fluorine interactions in left, cytosinium 5-fluorosaliclyate and right 5-fluorocytosinium salicylate molecular complexes.

The fluorine in the cytosinium 5-fluorosaliclyate molecular complex is involved in C-H \cdots F hydrogen bonds whilst the fluorine in the 5-fluorocytosinium salicylate molecular complexes is involved in both C-H \cdots F hydrogen bonds and fluorine-fluorine interactions. These interactions are what cause the subtle differences between the structures but have no effect over the primary bonding patterns that are adopted in all three structures. In the 5-fluorocytosine : 5-fluorosaliclyic acid hydrate molecular complex, the crystallisation ratio is different, and so are the primary hydrogen bonding motifs. There are both fluorine-fluorine interactions and C-H \cdots F hydrogen bonds in this molecular complex with the former acting as a link between separate layers whilst the weak hydrogen bond links chains within the same plane.

In the 5-fluorocytosine and 3-hydroxybenzoic acid molecular complex, the fluorine atom is involved in F \cdots F interactions help to link separate networks together. The main netlike structure is intertwined with another and the fluorine interactions are one of the main interactions present between these. The main interaction seen in this section is once again the C-H \cdots F interaction, however, the F \cdots F interactions play a prominent role in linking separate areas of the structure.

9.3.3 Disubstituted Benzoic Acid Molecular Complexes

C-H...F weak hydrogen bonds are again the most prominent fluorine interaction found in the molecular complexes of 5-fluorocytosine with disubstituted benzoic acids. The 5-fluorocytosinium 2,6-dihydroxybenzoate hydrate Form I molecular complex contains this hydrogen bond plays a significant role, with two of these interactions combining to form a base paired motif that helps to link heterodimers of hydrogen bonded rings between the two components (Figure 9.10).

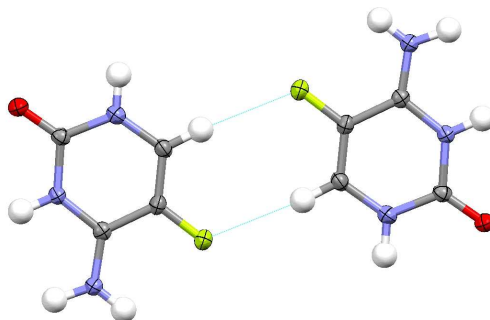


Figure 9.10. Base-pair style motif involving fluorine hydrogen bonds in the molecular complex of 5-fluorocytosinium 2,6-dihydroxybenzoate hydrate Form I.

In 5-fluorocytosinium 2,6-dihydroxybenzoate hydrate Form II, the fluorine does not form these interactions to form the base-paired motif, instead there are F...F interactions linking separate layers. However, there are still single C-H...F interactions but this time they are between 5-fluorocytosinium and 2,6-dihydroxybenzoate molecules linking separate heterodimers.

9.3.4 Thiourea and Urea Molecular Complexes

Only two molecular complexes were obtained with 5-fluorocytosine – 5-fluorocytosine and thiourea hydrate and 5-fluorocytosine and urea. In both of these molecular complexes, links are formed between chains via C-H...F hydrogen bonds (Figure 9.11). This bonding again resembles a base pairing motif between layers and once again shows that the C-H...F hydrogen bond is the most common bonding motif involving the fluorine atom.

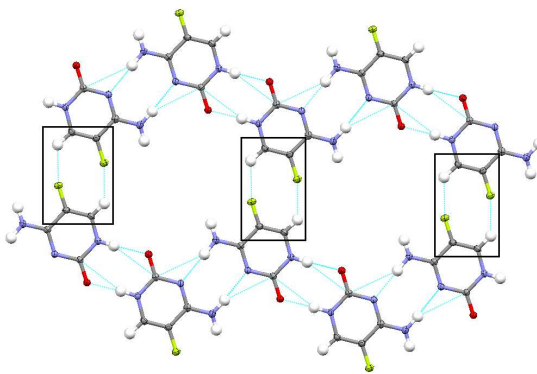


Figure 9.11. Fluorine interactions linking separate chains in the molecular complex of 5-fluorocytosine and urea.

9.3.5. Prevalence of Fluorine Interactions

In conclusion, the most common fluorine bond is the C-H...F hydrogen bond which often acts as the primary link between chains, but also as a link between molecules within a chain. In some cases, two C-H...F hydrogen bonds combine to form a base-pairing motif, however, this is rare with only four structures exhibiting this motif. Fluorine-fluorine interactions are not as common but when these have occurred they predominantly act as a link between layers. The strongest interactions in the molecular complexes are the base pairing motifs and the fluorine atoms do not disrupt these.

The length range of fluorine hydrogen bonds obtained in this thesis was 3.027(2)-3.595(2) Å with an average value of 3.361 Å. The length range for F...F interactions was 2.666(3)-2.801(2) Å with the average being 2.744 Å. The shortest of these are significantly shorter than the sum of the van der Waals radii for two fluorine atoms (2.94 Å); the longest are only just under the sum of the van der Waals radii for two fluorine atoms and thus may only be occurring as a consequence of other interactions rather than being a structure stabilising interaction.

9.4 5-Fluorouracil and Uracil Molecular Complexes

Three molecular complexes incorporating 5-fluorouracil and one containing uracil are reported. One of the most interesting results obtained was the pair of 5-fluorouracil and urea and uracil and urea molecular complexes. These two molecular complexes are isostructural with one another and the only difference between the two structures is the presence of the fluorine atom in one of the molecular complexes. There are two interactions in which the fluorine atom is involved. The first is a fluorine-fluorine interaction between two separate 5-fluorouracil molecules and the other interaction is an N-H...F hydrogen bond; this is the one of the few N-H...F hydrogen bond found in this

work. These interactions combine to have a subtle effect on the structure with the two slowly moving out of sync with one another as the two structures are overlaid and expanded (Figure 9.12). This moving out of sync is due to the interactions of the fluorine atom in the structure; these take the form of weak fluorine hydrogen bonds.

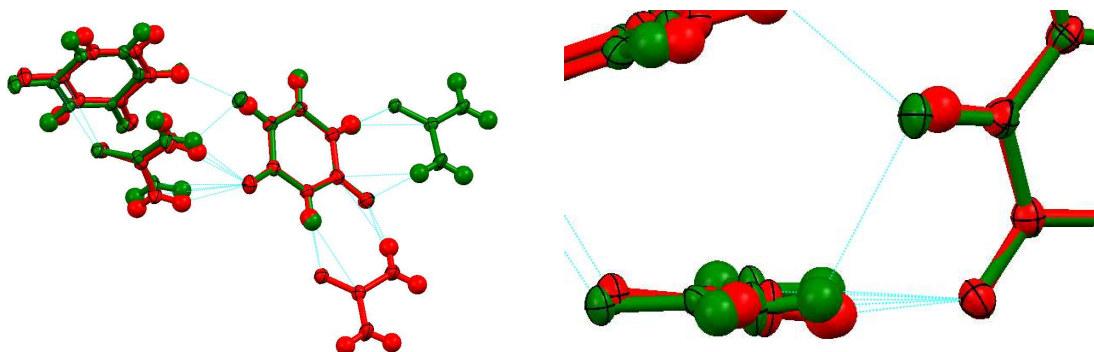


Figure 9.12. Overlaid molecular complexes of uracil and urea (red) and 5-fluorouracil and urea (green) showing subtle interactions due to the fluorine atom. Left shows the two structures moving out of sync and right shows the fluorine interactions which cause this slight change in orientation.

These two molecular complexes contain stacking interactions where pairs of uracil or 5-fluorouracil molecules are sandwiched by stacking interactions with the urea molecules. The stacking interactions between the uracil molecules are staggered meaning the uracil molecules do not line up perfectly. The orientation of the two urea and uracil molecules is also not the same (Figure 9.13).

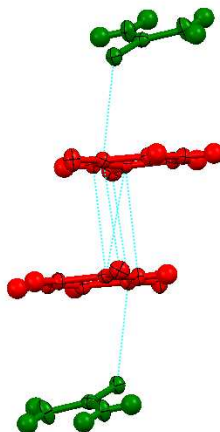


Figure 9.13. Stacking interactions in the uracil and urea molecular complex (red represents uracil and green represents urea)

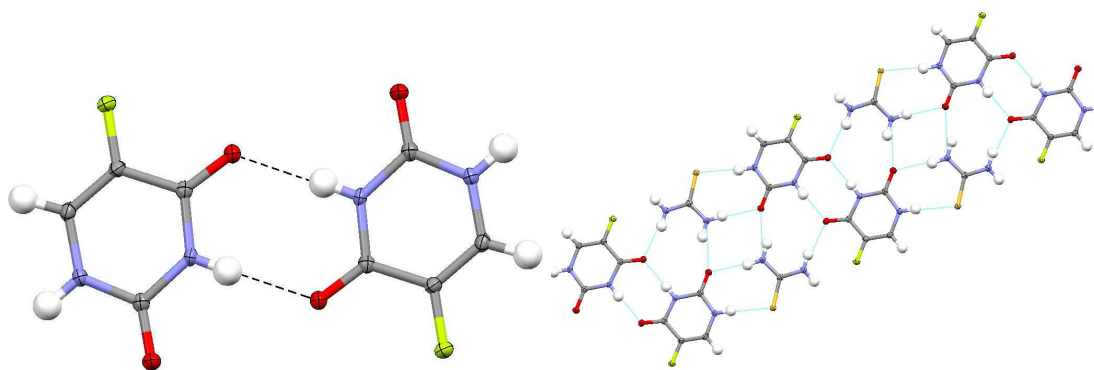


Figure 9.14. Homo base-pairing (left) and alternating pairs of 5-fluorouracil and thiourea molecules in the hydrogen bonded chain

The remaining two structures differ greatly most likely due to the inclusion of the larger sulphur atom. The 5-fluorouracil and thiourea structure contains homo-base-pairing (Figure 9.14, left) which is not seen in the other 5-fluorouracil molecular complexes. This homo-base-pairing, however, does not extend out into a homo-base paired chain; the thiourea molecule inserts into the chain to form a chain made up of a homo base pair followed by two thiourea molecules and this alternates as the chain is extended (Figure 9.14, right). This molecular complex also contains a fluorine hydrogen bond that involves the amine group of the thiourea molecule and is formed within the chain. There are stacking interactions in this molecular complex between two equivalent chains with one chain flipped so that the thiourea molecule stacks on top of the 5-fluorouracil molecule.

The last remaining molecular complex is the 1:4:5 5-fluorouracil and 3,5-dihydroxybenzoic acid hydrate. This structure contains a large number of water molecules. The most interesting interaction in the structure is the base paired motif comprised of two fluorine hydrogen bonds similar in nature the bonding motif observed in the 5-fluorocytosine and 3-fluorobenzoic acid molecular complex (and others). This is the only base pairing motif as this unit is surrounded by water molecules (Figure 9.15, top). These water molecules insert into the middle of the hydrogen bonds that would make up the normal homo-base-pairing motif (Figure 9.15, bottom). These water molecules act as a spacer in this structure. This structure contains stacking interactions but these are between the 3-5-dihydroxybenzoic acid molecules that surround a single base pair. This means that there is no possibility of base stacking interactions in the structure.

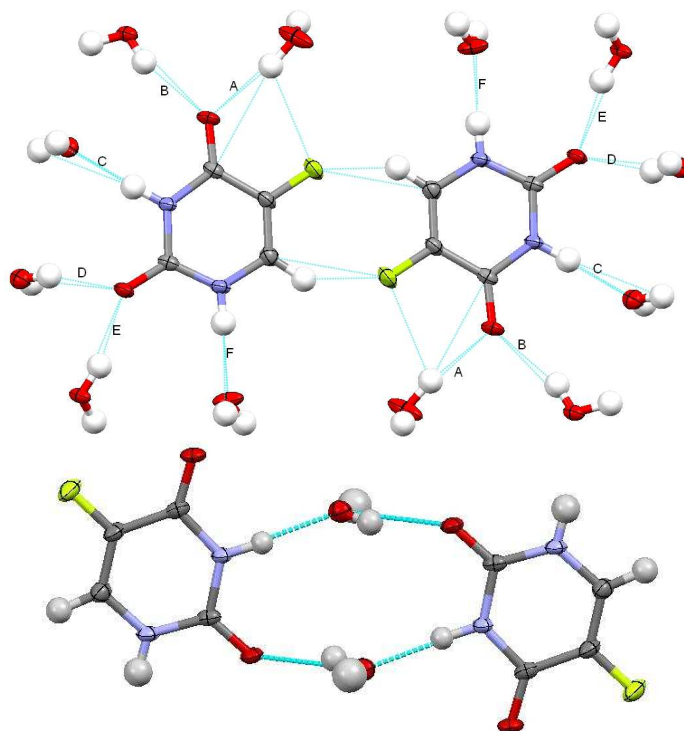


Figure 9.15. Top, base pairing motif in the molecular complex of 5-fluorouracil 3,5-dihydroxybenzoic acid hydrate involving C-H...F hydrogen bonds. Bottom, water molecules spacing the conventional base pair.

9.5 Stacking interactions

There are stacking interactions in the majority of the structures that have been presented in this thesis. The most common stacking interactions are pseudo-base-stacking interactions that are staggered where neutral cytosine molecules are stacked on top of charged cytosinium molecules. This type of stacking was very common and was seen, for example, in the molecular complex of cytosine and benzoic acid. When there are only neutral cytosine molecules present, the cytosine has been seen to have stacking interactions with other base molecules and also with the co-molecule present.

One noticeable theme that reoccurs throughout the structures is the fact that when there are extended base-paired chains, like those seen in the cytosine and benzoic acid structure, the charged cytosine molecule will stack on top of the neutral cytosine molecule. This is commonly seen in the molecular complexes that contain the pseudo-Watson-Crick hydrogen bonded chain. This is maintained even when the pseudo-Watson-Crick base pair does not result into a fully extended chain, in the example of 5-fluorocytosine and 4-hydroxybenzoic acid. The chain is not fully formed in this structure but once again the cytosinium stacks on top of the cytosine molecule.

The one structure that follows a different pattern from this set is the structure of cytosine and 3-fluorobenzoic acid. This structure forms the hydrogen-bonded ring motif and this

allows for stacking interactions present between 3-fluorobenzoate molecules that are sandwiched by cytosinium molecules via more stacking interactions. The majority of the structures that form the hydrogen-bonded ring motif contain a mixture of base stacking and normal stacking interactions as the pairs overlap. However, the related structures of cytosinium salicylate, 5-fluorocytosinium salicylate, and cytosinium 5-fluorosaliclate all form the hydrogen bonded ring. The stacking interactions are staggered to such an extent that the salicylate molecules stack on top of the cytosinium molecules and vice versa.

The molecular complexes involving the type A and type B extended chains demonstrate the variety of stacking interactions that can be obtained. The structures of 5-fluorocytosine with benzoic acid, 2-,3- and 4-fluorobenzoic acid all result in the same primary chain involving alternating type A and type B bonding motifs. However, the benzoic acid structure and the 2-fluorobenzoic acid molecular complexes have stacked chains flipped with respect to one another. However, the structures with 3- and 4-fluorobenzoic acid have a simple staggered stacked chain where the cytosine molecules are stacked and orientated in the same direction.

In conclusion, when there are extended pseudo-Watson-Crick base paired chains it is a common occurrence for the cytosinium molecule to stack on top of the cytosine molecule. Where there are only charged molecules in the molecular complexes and the hydrogen bonded heterodimers are formed, cytosinium molecules tend to stack on top of the co-molecules. This suggests that the stacking interactions are influence by the presence of charge. If the type A and type B extended chain is formed a greater variety of stacking is obtained.

9.6 dSNAP

dSNAP has been used to classify the preferred hydrogen bonding motifs of the natural base pairs contained within the CSD. However, the conclusions that could be drawn were limited by the relatively few structures of this type found. The clustering of adenine and thymine worked well. This search and subsequent clustering gave 11 clusters with three preferred bonding motifs, the Watson Crick, the pseudo Watson Crick and the Hoogsteen. These were represented by 8, 7 and 8 hits, respectively.

When carrying out a similar analysis on cytosine with guanine, fewer hits were obtained and this might be explained by solubility problems, notably with guanine. These structures were found to be best classified by three clusters. The preferred hydrogen bonding motif is the standard Watson Crick motif.

The dSNAP program was successful in separating out the hits due to the prominent base-pairing bonding motifs adopted by the specific hit, however, the program did not give an insight into the buckling or the propeller twisting in the structures. Different values of propeller twisting and buckling was found to appear in several different clusters and separate into different clusters.

9.7 Further Work

One of the most interesting sets of molecular complexes obtained in this thesis are the 3-fluorobenzoic acid and 4-fluorobenzoic acid with 5-fluorocytosine complexes. These structures both contain channels that are formed via fluorine hydrogen bonds. These channels contain disordered solvent and from the results obtained from TGA and DSC analyses, the possibility of removing the solvent before the structure was destroyed was identified. Further investigations are required to try and remove the solvent and see if the crystal can be recovered and the structure determined in the absence of solvent in the channels. If this process is possible, then it may be possible to fill these channels with gases such as H₂ or CO₂ and apply these as gas storage materials.

dSNAP has been shown to provide a useful method for classifying the main supramolecular units in nucleobase molecular complexes. This could be expanded to include other molecular complexes in addition to the original hits; as a first step, the molecular complexes obtained within this thesis will be added to those already reported in the CSD to enable a comparison across a wider range of materials.

This work could be extended to include fluorinated nucleosides and nucleotides and the effect of the positioning of the fluorine on their function. This could involve fluorinating the base, sugar or phosphate group. This would lead to further understanding of any structural effects that may be relevant to the use of these materials for potential applications in the treatment of cancer and viral cells due to the different effect that the presence of fluorine will have on the efficacy of these.

This work has shown that fluorination can have a significant effect on some of the interactions of importance in DNA structure, notably base-pairing, base-stacking and also by the introduction of additional intermolecular interactions in the form of fluorine interactions. Future extensions of this work could therefore also include an investigation into the possible effect of these modified interactions on the self-assembly and stability of fluorinated DNA double helical structures.

10. References

1. J. S. Lai, E. T. Kool, *Chem. Eur. J.*, 2005, **11**, 2966
2. E.T, Kool, *Curr Opin Chem Biol* 4, 2000, 802
3. E.T. Kool, *Acc Chem Res*, 2002, **325**, 936
4. J.S. Lai, E.T. Kool, *Chem Eur J*, 2005, **11**, 2966
5. A. Parkin, S. M. Harte, A. E. Goeta, C. C. Wilson, *New J. Chem.*, 2004, **28**, 718
6. S. M. Harte, A. Parkin, A. Goeta, C. C. Wilson, *Journal of Molecular Structure*, 2005, **741**, 93
7. A. Parkin, M. Adam, S. M. Harte, J. L. Kennedy, C. C. Wilson, *Acta Cryst.*, 2006, **E62**, 987
8. J. D. Watson, F. H. C. Crick, *Nature* 2005, **171**, 737
9. <http://www.accessexcellence.org/RC/VL/GG/dna2.php> (last accessed 2011-09-12)
10. P. Yakovchuk, E. Protozanova, M. D. Frank-Kamenetskii, *Nucleic Acids Research*, 2006, **34**, 564
11. V. L. Damaraju, S. Damaraju, J. D. Young, S. A. Baldwin, J. Mackey, M. B. Sawyer, C. E. Cass, *Oncogene*, 2003, **22**, 7524
12. T. Zhan; H. Lou, *Yaoxue Jinzhan*, 2005, **29**(7), 289
13. M. Ji, *Zhongguo Yaoshi*, 2006, **9**(11), 1054
14. S. A. Thomas, *Current Pharmaceutical Design*, 2004, **10**(12), 1313
15. V. L. Damaraju, S. Damaraju, J. D. Young, S. A. Baldwin, J. Mackey, M. B. Sawyer, C. E. Cass. *Oncogene*, 2003, **22**(47), 7340
16. W. Meng, F. Qing, *Current Topics in Medicinal Chemistry*, 2006, **6**(14), 1499
17. K. W. Pankiewicz, *Carbohydrate Research*, 2000, **327**(1-2), 87
18. H. Kim, S.G. DiMagno, Book of Abstracts, 213th ACS National Meeting, San Francisco, April 13-17 (1997), FLUO-002.
19. J. Chang, X. Yu, X. Bao, Z. Ye, *Faming Zhuanli Shenqing Gongkai Shuomingshu*, 2005, 12.
20. J. F. C. Turner, University of Tennessee, Knoxville, Tennessee 37996-1600
21. H. Ghazi, E. Lee-Ruff, *Journal of Fluorine Chemistry*, 2005, **126**(11-12), 1565
22. M.L Waters, *Current Opinion in Chemical Biology*, 2002, **6**(6), 736
23. J.Y. Douillard, D. Cunningham, A.D. Roth, *Lancet*, 2000, **355**, 1041
24. S. Giacchetti, B. Perpoint, R. Zidani, N. Le Bail, R. Faggiuolo, C. Focan, P. Chollet, J.F. Llory, Y. Letourneau, B. Coudert, F. Bertheaut-Cvitkovic, D. Larregain-Fournier, A. Le Rol, S. Walter, R. Adam, J.L. Misset, and F. Le'vi, *J. Clin. Oncol.*, 2000, **18**, 136

25. D.B. Longley, P. Harkin, P. G. Johnston, *Nature Reviews Cancer* **3**, 330
26. K. Park, N. R. Kitteringham, *Annu. Rev. Pharmacol. Toxicol.* 2001, **41**, 443
27. M. J. Zaworotko, *Nature Chemistry*, 2011,**3**(9),653
28. P. Erk, H. Hengelsberg, M. F. Haddow, R. van Gelder, *CrystEngComm*, 2004, **6**, 474
29. C.B. Aakeroy, N.R. Champness, C.Janiak, *CrystEngComm*, 2010, **12**, 22
30. M. S. Adam, A. Parkin, L.H. Thomas, C.C. Wilson, *CrystEngComm*, 2010, **12**, 917
31. An Introduction to Hydrogen Bonding (Jeffrey, George A.), Michael J. Minch
Department of Chemistry, University of the Pacific, Stockton, CA 95211
32. J.J. Benítez, J.A. Heredia-Guerrero, A.Heredia, *J. Phys. Chem. C*, 2007, **111** (26), 9465
33. G.Tóth, C.R. Watts, R.F. Murphy, S.L. Proteins: *Structure, Function, and Genetics*, 2001, **43**(4), 373
34. R. West, D.L. Powell, L.S. Whatley, M.K.T. Lee, P.V. Rague Schleyer, *J. Am. Chem. Soc.*, 1962,**84**, 3221
35. D. O'Hagan, H.S. Rzepa, *Chem. Commun.* 1997, 645
36. P. Murray Rust, W.C. Stallings, C.T. Monti, R.K. Preston, J.P. Glusker, *J. Am. Chem. Soc.* 1983, **105**, 3206
37. L. Brammer, E.A. Bruton, P. Sherwood, *Cryst. Growth Des.* 2001, **1**, 277
38. Q. Yao, A.R. Motta, *Tetrahedron Lett.*, 2004, **45**, 777
39. F. H. Allen, *Acta Cryst.*, 2002, **B58**, 380
40. L. Young, V.V. Prabhu, E.W. Prohofsky, *Phys. Rev.*, 1989, A **40**, 5451
41. P. A. Giguere, *J. Chem. Phys.* 1987, **87**, 4835
42. J. Sponer, and P. Hobza, *J. Am. Chem. Soc.* 1994, **116**, 709
43. P. Gilli, V. Bertolasi, V. Ferretti, *J. Am. Chem. Soc.* 1994, **116**, 909
44. P. Gilli, V. Bertolasi, L. Pretto, V. Ferretti and G. Gilli, *J. Am. Chem. Soc.* 2004,**126**, 3845
45. P. Gilli, V. Bertolasi, L. Pretto, and G. Gilli, *J. Mole. Struct.* 2006, **790**, 40
46. T. Steiner, I. Majerz, and C. C. Wilson, *Angew. Chem. Int. Ed.* 2001, **40**, 2651
47. Y. Enomoto, G. Honda, Y. Shimizu, S. Matsui,. M. Sakata, O. O. Drozdova, K. Yakushi, *Bull.Chem.Soc.Jpn.*, 2008, **81**, 331
48. G.A. Sim, J.M. Robertson, T.H. Goodwin *Acta Cryst.*, 1955, **8**, 157
49. S.L. Childs, G.P. Stahly, A. Park, *Mol. Pharmaceuticals*, 2007, **4** (3), 323
50. J.A. Gerlt, P.G. Gassman, *Biochemistry*, 1993, **32** (45), 11943
51. B. Schwartz, D.G. Drueckhammer, K.C. Usher, S.J. Remington, *Biochemistry* 1995, **34**, 15459

52. M. Schmidtman, C.C. Wilson, *CrystEngComm*, 2008, **10**, 177
53. G.R. Desiraju, R. Parthasarathy, *J. Am. Chem. Soc.*, 1989, **111** (23), 8725
54. J.D. Dunitz, *Philos. Trans. R. SOC. London*, 1975, **5272**, 99
55. D. E. Williams, L. Y Hsu, *Acta Crystallogr.*, 1985, **A41**, 296
56. T. Sakurai, M.Sundaralingam, G.A. Jeffrey, *Acta Crystallogr.*, 1963, **16**, 354
57. S. C. Nyburg,.; W.Wong-Ng, *Proc. R. SOC. London*, 1979, **A367**, 29
58. Wei-Qun Zhou, Lu-De Lu, Xu-Jie Yang, Yang Cao, *Acta Chim.Sinica*, 2004, **62**, 1472
59. G.A. Jeffrey, Y Kinoshita, *Acta Cryst.*, 1963, **16**, 20
60. S. Neidle, A. Achari, M. Rabinovitch, *Acta Cryst.*, 1976, **B32**, 2050
61. M. Eisenstein, *Acta Cryst.*, 1988, **B44**, 412
62. R.J. McClure, B.M. Craven, *Acta Cryst.*, 1973, **B29**, 1234
63. O. Munshi, T.N. Guru Row, *Acta Cryst.*, 2006, **B62**, 612
64. D.L. Barker, R.E. Marsh, *Acta Cryst.*, 1964, **17**, 1581
65. K. Bouchouit, N. Benali-Cherif, S. Dahaoui, E.E. Bendeif, C. Lecomte, *Acta Cryst.*, 2005, **E61**, 2755
66. T.R. Kulkarni, *J.Mol.Struct.*, 2008, **873**, 160
67. D. Voet, A. Rich, *J.Am.Chem.Soc.*, 1969, **91**, 3069
68. G. Portalone, M. Colapietro, *Acta Cryst.*, 2007, **C63**, 423
69. C. Tamura, S. Sato, T. Hata, *Bull.Chem.Soc.Jpn.*, 1973, **46**, 2388
70. G. Portalone, M. Colapietro, *J.Chem.Cryst.*, 2009, **39**, 193
71. N. Benali-Cherif, W. Falek, A. Direm, *Acta Cryst.*, 2009, **E65**, 3058
72. B. Sridhar, K.Ravikumar, *Acta Cryst.*, 2008, **C64**, 566
73. S. Sathiya Moorthi, K. Chinnakali, S. Nanjundan, P. Selvam, H.-K. Fun and X.-L. Yu *Acta Cryst.*, 2005, **E61**, 743
74. T. Balasubramanian, P.T. Muthiah, W.T. Robinson, *Bull. Chem. Soc. Jpn*, 1996, **69**, 2919
75. M. Gdaniec, B. Brycki, M. Szafran, *J.Mol.Struct.*, 1989, **195**, 57
76. R. Thomas, G.U. Kulkarni., *Journal of Molecular Structure*, 2008, **873**, 160
77. S.R. Perumalla, E. Suresh, V.R, Pedireddi, *Angew.Chem.,Int.Ed.*, 2005, **44**, 7752
78. A. Vermes, H.-J. Guchelaar, H.-J. Guchelaar, *J. Antimicrob. Chemother.*, 2000, **46**(2), 171
79. T. Louis, J.N. Low, P. Tollin, *Cryst.Struct.Comm.*, 1982, **11**, 1059
80. A.T. Hulme, D.A. Tocher, *Crystal Growth & Design*, 2006, **6** (2), 481
81. P.Prabakaran, S. Murugesan, P.T. Muthiah, G. Bocelli, L. Righi, *Acta Cryst.*, 2001, **E57**, 933

82. R. Gerdil, *Acta Cryst.*, 1961, **14**, 333
83. Z. Li, J.F. Ding, G. Robertson, M. Day, Y. Tao, *Tetrahedron Lett.*, 2005, **46**, 6499
84. T. Sakurai, M. Okunuki, *Acta Cryst.*, 1971, **B27**, 1445
85. S.M. Tret'yak, V.V. Mitkevich, L.F. Sukhodub,
Kristallografiya(Russ.)(Crystallogr.Rep.), 1987, **32**, 1268
86. B. Sridhar, K. Ravikumar, *Acta Cryst.*, 2007, **C63**, 212
87. S. Mahapatra, S.K. Nayak, S.J. Prathapa, T.N. Guru Row, *Cryst.Growth Des.*,
2008, **8**, 1223
88. U. Thewalt, C.E. Bugg, R.E.Marsh, *Acta Cryst.*, 1971. **B27**, 2358
89. C.E. Bugg, U. Thewalt, *Acta Cryst.*, 1975, **B31**, 121
90. R.F. Stewart, L.H. Jensen, *Acta Cryst.*, 1967, **23**, 1102
91. R. Thomas, G.U. Kulkarni, *Beilstein J.Org.Chem.*, 2007, **3**, 17
92. B.J Cummings, T.J.Keane, B.O'Sullivan, C.S. Wong, C.N. Catton, *J. Radiorron
Oncology Biol. Pkys.*, 1991, **21**, 1115
93. J.M. Collins, *Clin Cancer Res*, 2000, **6**, 1322
94. L. Fallon III, *Acta Cryst.*, 1973, **B29**, 2549
95. A.T. Hulme, S.L. Price, D.A. Tocher, *J Am Chem Soc*, 2005, **127**(4), 1116
96. G. Portalone, M. Colapietro, *J.Am.Chem.Soc.*, 1969, **91**, 3069
97. S. R. Perumalla, E. Suresh, V.R. Pedireddi, *Angew. Chem. Int. Ed.*, 2005, **44**, 7752
98. D.Armentano, G.De Munno, R.Rossi, *New J.Chem.*, 2006, **30**, 13,
99. P.-Z. Hu, L.-F. Ma, J.-G. Wang, B.-T. Zhao, L.-Y. Wang, *Acta Cryst.*, 2005, **E61**,
2775
100. K.Bouchouit, N.Benali-Cherif, S.Dahaoui, E.-E.Bendeif, C.Lecomte, *Acta Cryst.*
2005, **61**, 2755, 65
101. P.E. Bourne, M.R. Taylor, *Acta Cryst.* (1983). **C39**, 430
102. C.C. Wilson, P. Tollin, *Nucleosides and Nucleotides*, 1987, **6**, 643
103. C.C. Wilson, *Nucl Acids Res*, 1987, **15**, 8577
104. C.C. Wilson, *Nucleosides and Nucleotides*, 1990, **9:4**, 479
105. P. Hobza, C. Sandorfy, *J. Am. Chem. Soc.*, 1987, **109**, 1302
106. A.D. Di Liscia, "Johannes Kepler", *The Stanford Encyclopedia of Philosophy
(Summer 2011 Edition)*, Edward N. Zalta (ed.),
107. J G Crowther, *Founders of British science : John Wilkins, Robert Boyle, John
Ray, Christopher Wren, Robert Hooke, Isaac Newton*,
108. J.Z. Buchwald, *Arch. Hist. Exact Sci.*, 2007, **61**, 67
109. Nicolaus Steno (Niels Stensen) and Erasmus Bartholinus;: Two 17th century
Danish scientists and the foundation of exact geology and crystallography

110. Essai de Cristallographie: *Ou Description Des Figures Geometriques* (1772),
Jean Baptiste Louis De Rome De L'Isle
111. J. G. Burke, *Origins of the Science of Crystals* (Berkeley-Los Angeles, 1966)
112. W.H. Miller, *A Treatise on Crystallography*, 1839
113. W.L. Bragg, (1913). "The Diffraction of Short Electromagnetic Waves by a
Crystal". *Proceedings of the Cambridge Philosophical Society* 17: 43
114. L. V. Azároff, R. Kaplow, N. Kato, R. J. Weiss, A. J. C. Wilson, R. A. Young, X-
ray diffraction, 1974, McGraw-Hill.
115. O. Glatter; O. Kratky, *Small Angle X-ray Scattering*, 1982. Academic Press
116. L. J. Farrugia, *J. Appl. Cryst.*, 1999, **32**, 837
117. P. W. Betteridge, J. R. Carruthers, R. I. Cooper, K. Prout, D. J. J. Watkin, *Appl.*
Cryst., 2003, **36**, 1487
118. M. A. Cooper, F. C. Hawthorne, *Can. Mineral*, 1995, **33**, 1125
119. CrystalClear: An Integrated Program for the Collection and Processing of Area
Detector Data, Rigaku Corporation, © 1997-2002
120. Otwinowski, Z. & Minor, W., *Methods in Enzymology*, 1997, **276**, 307
121. G.Barr, W. Dong, C.J. Gilmore, *J. Appl. Cryst.*, 2009, **42**, 965
122. R. J. Hill, C. J. Howard, *J. Appl. Cryst.*, 1987, **20**, 467
123. G. He, V. Bhamidi, S.R. Wilson, R.B.H. Tan, P.J.A. Kenis, C.F. Zukoski, *Crystal*
Growth & Design, 2006, **6** (8), 1746
124. N. E. Chayen, *Acta Cryst.*, 1998, **D54**, 8
125. M.Ercken, P.Adriaenssens, D.Vanderzande, Jan Gelan *Macromolecules*, 1995,
28(25), 8541
126. [http://thermalguides.com/thermogravimetric-analysis-for-compositional-analysis-
of-multi-component-polymers/](http://thermalguides.com/thermogravimetric-analysis-for-compositional-analysis-of-multi-component-polymers/) (last accessed 13/SEPT/2011)
127. G.Barr, W.Dong, C.J. Gilmore, A. Parkin, C.C. Wilson, *Appl. Cryst.*, 2005, **38**,
833
128. S.R.Perumalla, E.Suresh, V.R.Pedireddi, *Angew.Chem.,Int.Ed.*, 2005, 44, 7752
129. D.L. Barker, R.E. Marsh, *Acts Cryst.*, 1964, **17**, 1561
130. An Introduction to Hydrogen Bonding (Jeffrey, George A.), Michael J. Minch
Department of Chemistry, University of the Pacific, Stockton, CA 95211
131. E.J.O'Brien, *Acta Crystallogr.*, 1967, **23**, 92
132. http://web.mit.edu/platon_v40505/platon/docs/platon/pl000303.html (last
accessed 2nd April 2011)
133. Harte. Suzanne Marie, Variable temperature diffraction studies of hydrogen
bonded dimer systems, Thesis (Ph.D.) - University of Glasgow, 2006.

134. P. Prabakaran, S. Murugesan, P. T. Muthiah, G. Bocelli, L. Righi, *Acta Cryst*, 2001, **E57**, 933
135. B.R. Bhogala, S. Basavoju, A. Nangia, *CrystEngComm*, 2005, **7**, 551
136. M.Colapietro, A.Domenicano, C.Marciante, G.Portalone, *Eur.Cryst.Meeting*, 1983, **8**, 124
137. I. J. Bruno, J. C. Cole, P. R. Edgington, M. Kessler, C. F. Macrae, P. McCabe, J. Pearson, R.Taylor, *Acta Cryst.*, 2002, **B58**, 389
138. A. Parkin, G. Barr, W. Dong, C. J. Gilmore, C. C. Wilson, *CrystEngComm*, 2006,**8**, 257
139. A. Collins, A. Parkin, G. Barr, W. Dong, C. J. Gilmore, C. C. Wilson, *CrystEngComm*, 2007,**9**, 245
140. P.Hobza, C.Sandorfy, *J. Am. Chem. Soc.*, 1987, **109** (5), 1302
141. C. F. Guerra, F. M. Bickelhaupt, J. G. Snijders and E. J. Baerends, *J. Am. Chem. Soc.* 2000, **122**, 4117
142. S.Fujii, K.Kawasaki, A.Sato, T.Fujiwara, K.-I.Tomita, *Arch.Biochem.Biophys*, 1977, **181**, 363.
143. M.Salas, B.Gordillo, F.J.Gonzalez, *ARKIVOC*, 2003, **4**, 72
144. T.D.Sakore, H.M.Sobell, F.Mazza, G.Kartha, *J.Mol.Biol.*, 1969, **43**, 385
145. D. Voet, Alexander Rich, *J. Am. Chem. Soc.*, 1972, **94** (16), 5888
146. D. Voet, *J. Am. Chem. Soc.*, 1972, **94** (23), 8213
147. A. Camerman, D. Mastropaolo, A. Hempel, N. Camerman, *Canadian Journal of Chemistry*, 2000, **78**, 1045
148. K.Tomita, L.Katz, A.Rich, *J.Mol.Biol.*, 1967, **30**, 545
149. K. Hoogsteen, *Acta Cryst.*, 1963, **16**, 907
150. T.D.Sakore, S.S.Tavale, H.M.Sobell, *J.Mol.Biol.*, 1969, **43**, 361
151. T.D.Sakore, H.M.Sobell, F.Mazza, G.Kartha, *J.Mol.Biol.*, 1969, **43**, 385
152. I.-N. Hsu and B. M. Craven, *Acta Cryst.*, 1974, **B30**, 988
153. S.H. Kim, A. Rich., *Proc. Natl. Acad. Sci.*, 1968. **60**, 402
154. X. Zhang, B.Bernet, A.Vasella, *Helv.Chim.Acta*, 2006, **89**, 2861
155. G.Simundza, T.D.Sakore, H.M.Sobell, *J.Mol.Biol.*, 1970, **48**, 263
156. T. Ise, D. Shiomi, K. Sato, T. Takui, *Chem. Commun.*, 2006, 4832
157. T.Ishida, Y.Tokura, M.Shimamoto, M.Doi, M.Inoue, *Chem.Pharm.Bull.*, 1987, **35**, 1691
158. S.S.Tavale, T.D.Sakore, H.M.Sobell, *J.Mol.Biol.*, 1969, **43**, 375
159. G.Simundza, T.D.Sakore, H.M.Sobell, *J.Mol.Biol.*, 1970, **48**, 263
160. S.Fujita, A.Takenaka, Y.Sasada, *Bull.Chem.Soc.Jpn.*, 1984, **57**, 1707

161. S. Fujita, A. Takenaka, Y. Sasada, *Biochemistry*, 1985, **24** (2), 508
162. H. Tanaka, D. Shiomi, T. Ise, K. Sato, T. Takui, *CrystEngComm*, 2007, **9**, 767
163. T. Murata, K. Nishimura, G. Saito, *Mol. Cryst. Liq. Cryst. Sci. Technol., Sect. A*, 2007, **466**, 101
164. S. R. Perumalla, E. Suresh, V. R. Pedireddi, *Angew. Chem., Int. Ed.*, 2005, **44**, 7752
165. T. Murata, G. Saito, K. Nishimura, Y. Enomoto, G. Honda, Y. Shimizu, S. Matsui, M. Sakata, O. O. Drozdova, K. Yakushi, *Bull. Chem. Soc. Jpn.*, 2008, **81**, 331
166. Jiabao Weng, Maochun Hong, Yucang Liang, Qian Shi and Rong Cao, *J. Chem. Soc., Dalton Trans.*, 2002, 289
167. K. Ogawa, M. Kumihashi, K. Tomita, S. Shirotake, *Acta Cryst.*, 1980, **B36**, 1793
168. P. Hobza, C. Sandorfy, *J. Am. Chem. Soc.*, 1987, **109** (5), 1302
169. K. Tamura, Y. Ogo and T. Imoto, *Bull. Chem. Soc. Jpn.*, 1973, **46**, 2388
170. G. A. Jeffrey and Y. Kinoshita, *Acta Crystallogr.*, 1963, **16**, 20
171. K. C. Kumara Swamy, S. Kumaraswamy, P. Kommana, *J. Am. Chem. Soc.*, 2001, **123**, 12642
172. I. J. Bruno, J. C. Cole, P. R. Edgington, M. Kessler, C. F. Macrae, P. McCabe, J. Pearson, R. Taylor, *Acta Cryst.*, 2002, **B58**, 389
173. K. Bouchouit, N. Benali-Cherif, S. Dahaoui, E.-E. Bendeif, C. Lecomte, *Acta Crystallogr., Sect. E: Struct. Rep. Online*, 2005, **61**, 2755,
174. F. Fujinami, K. Ogawa, K. Arakawa, S. Shirotake, S. Fujii, K. Tomita, *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.*, 1979, **35**, 968,
175. <http://www.theory.nipne.ro/~dragos/Solid/FYS230-Exercises.html> (last accessed 19th Jan 2012)
176. G. M. Sheldrick, *Acta Cryst.*, 2008, **A64**, 112

11. Appendix

11.1 Adenine and Thymine dSNAP study

11.1.1 Yellow and Red Cluster

ADRBFT10 - S.Fujii, K.Kawasaki, A.Sato, T.Fujiwara, K.-I.Tomita, *Arch.Biochem.Biophys.*, 1977, 181, 363.

11.1.2 Green Cluster

KECYUN - M.Salas, B.Gordillo, F.J.Gonzalez, *ARKIVOC*, 2003, 4, 72

MTYDAP - T.D.Sakore, H.M.Sobell, F.Mazza, G.Kartha, *J.Mol.Biol.*, 1969, 43, 385

11.1.3 Cyan Cluster

EADBAC – D.Voet, A.Rich, *J. Am. Chem. Soc.*, 1972, 94 (16), 5888

EADBAR10 – D.Voet, *J. Am. Chem. Soc.*, 1972, 94 (23), 8213

QEDFEK – A.Camerman, D.Mastropaolo, A.Hempel, N.Camerman, *Canadian Journal of Chemistry*, 2000, 78:1045

ETABFU - K.Tomita, L.Katz, A.Rich, *J.Mol.Biol.*, 1967, 30, 545

MTHMAD - K. Hoogsteen, *Acta Cryst.*, 1963, 16, 907

IMUEAD - T.D.Sakore, S.S.Tavale, H.M.Sobell, *J.Mol.Biol.*, 1969, 43, 361

MIUDAP10 - T.D.Sakore, H.M.Sobell, F.Mazza, G.Kartha, *J.Mol.Biol.*, 1969, 43, 385

MBUMAD10 - I.-N. Hsu, B. M. Craven, *Acta Cryst.*, 1974, B30, 988

11.1.4 Blue Cluster

BARAAD - S.H. Kim, A. Rich. *Proc. Natl. Acad. Sci. U.S.A.*, 1968, 60, 402.

11.1.5 Pink Cluster

ETABFU - K.Tomita, L.Katz, A.Rich, *J.Mol.Biol.*, 1967, 30, 545

MTHMAD - K. Hoogsteen, *Acta Cryst.*, 1963, 16, 907

11.1.6 Orange Striped

- BARAAD - S.H. Kim, A. Rich. *Proc. Natl. Acad. Sci. U.S.A*, 1968, 60, 402.
- EADBAR10 – D.Voet, *J. Am. Chem. Soc.*, 1972, 94 (23), 8213
- QEDFEK – A.Camerman, D.Mastropaolo, A.Hempel, N.Camerman, *Canadian Journal of Chemistry*, 2000, 78, 1045
- EADBAC – D.Voet, A.Rich, *J. Am. Chem. Soc.*, 1972, 94 (16), 5888
- DETVAA_01 – X. Zhang, B.Bernet, A.Vasella, *Helvetica Chimica Acta*, 2006, 89, 2861
- DETVAA_02 - X. Zhang, B.Bernet, A.Vasella, *Helvetica Chimica Acta*, 2006, 89, 2861
- MTYDAP - T.D.Sakore, H.M.Sobell, F.Mazza, G.Kartha, *J.Mol.Biol.*, 1969, 43, 385
- EBPMUR – G.Simundza, T. D. Sakore, Henry M. Sobell, *Journal of Molecular Biology*, 1970, 48(2), 263

11.1.7 Green Striped

- BARAAD - S.H. Kim, A. Rich. *Proc. Natl. Acad. Sci. U.S.A*, 1968, 60, 402
- BURBAD - S.S.Tavale, T.D.Sakore, H.M.Sobell, *J.Mol.Biol.*, 1969, 43, 375
- CETQUO – T. Ise, D. Shiomi, K. Sato, T. Takui, *Chem. Commun.*, 2006, 4832
- FUJDET - T.Ishida, Y.Tokura, M.Shimamoto, M.Doi, M.Inoue, *Chem.Pharm.Bull.*, 1987, 35, 1691
- IMUEAD - T.D.Sakore, S.S.Tavale, H.M.Sobell, *J.Mol.Biol.*, 1969, 43, 361
- MIUDAP10 - T.D.Sakore, H.M.Sobell, F.Mazza, G.Kartha, *J.Mol.Biol.*, 1969, 43, 385
- EBPETY - G.Simundza, T.D.Sakore, H.M.Sobell, *J.Mol.Biol.*, 1970, 48, 263

11.1.8 Cyan Striped

- KECYUN - M.Salas, B.Gordillo, F.J.Gonzalez, *ARKIVOC*, 2003, 4, 72

11.1.9 Blue Striped

- FUJDET - T.Ishida, Y.Tokura, M.Shimamoto, M.Doi, M.Inoue, *Chem.Pharm.Bull.*, 1987, 35, 1691

11.1.10 Purple Striped

- MIUDAP10 - T.D.Sakore, H.M.Sobell, F.Mazza, G.Kartha, *J.Mol.Biol.*, 1969, 43, 385

11.2 Guanine and Cytosine dSNAP study

11.2.1 Red Cluster

BUDWAY - S.Fujita, A.Takenaka, Y.Sasada, *Bull.Chem.Soc.Jpn.*, 1984, 57, 1707

EGMCYT – S.Fujita, A.Takenaka, Y.Sasada, *Biochemistry*, 1985, 24 (2), 508

MILFAP - H.Tanaka, D.Shiomi, T.Ise, K.Sato, T.Takui, *CrystEngComm*, 2007, 9, 767

DIBMEG - H.Tanaka, D.Shiomi, T.Ise, K.Sato, T.Takui, *CrystEngComm*, 2007, 9, 767

EGMFCY - E.J.O'Brien, *Acta Crystallogr.*, 1967, 23, 92

11.2.2 Yellow Cluster

MILFAP - H.Tanaka, D.Shiomi, T.Ise, K.Sato, T.Takui, *CrystEngComm*, 2007, 9, 767

11.2.3 Green Cluster

BUDWAY - S.Fujita, A.Takenaka, Y.Sasada, *Bull.Chem.Soc.Jpn.*, 1984, 57, 1707

11.2.4 Cyan Cluster

EGMCYT – S.Fujita, A.Takenaka, Y.Sasada, *Biochemistry*, 1985, 24 (2), 508

MILFAP - H.Tanaka, D.Shiomi, T.Ise, K.Sato, T.Takui, *CrystEngComm*, 2007, 9, 767